

---

# REPORT

---

---

## **An Evaluation of the National Electronic Telecommunications System for Surveillance (NETSS)**

---

---

To

---

The Centers for Disease Control

---

Office of Program Planning and

---

Evaluation and

---

the Epidemiology Program Office

---

Atlanta, Georgia

---

---

June 30, 1991

---



**Battelle**

*Putting Technology To Work*

Arlington Office  
2101 Wilson Boulevard  
Suite 800  
Arlington VA 22201-3008  
Telephone (703) 528-6603  
Facsimile (703) 527-5640

June 30, 1991

Ms. Floy Cross , Project Officer  
Office of Program Planning and Evaluation  
Centers for Disease Control  
1600 Clifton Road  
Mailstop D24  
Atlanta, GA 30333

Dear Floy:

**Project No. 200-88-0642**

Enclosed are six (6) copies the Final Report for Task 8, "An Evaluation of the National Electronic Telecommunications System for Surveillance (NETSS)", conducted under the above contract number. This document is specified as a deliverable in the Scope of Work.

If you have any questions, please call me at (703) 875-2966 or Bruce Ellis (703) 875-2958.

Sincerely,

Mary E. Odell-Butler, PhD  
Principal Research Scientist  
Statistics and Data Analysis Systems

Bruce Ellis, M.S.  
Research Scientist  
Statistics and Data Analysis Systems

MEO/jbs

cc: Nancy Chalmers, EPO  
Barbara Panter-Connah, EPO  
Bradford Myers, CDC contracts (Letter Only)

**June 30, 1991**

**FINAL REPORT**

**for**

**TASK 8**

**AN EVALUATION OF THE NATIONAL ELECTRONIC  
TELECOMMUNICATIONS SYSTEM FOR SURVEILLANCE (NETSS)**

**Mary E. Odell-Butler, Ph.D.  
Task Leader**

**Bruce Ellis, M.S.**

**James C. Hersey, Ph.D.  
Project Director**

**BATTELLE  
Statistics and Data Analysis Systems  
210 1 Wilson Boulevard  
Arlington, Virginia 22201**

**Contract Number 200-88-0642**

**Floy Cross, Project Officer  
Barbara Panter-Connah, Technical Monitor  
Nancy Chalmers, Technical Monitor**

**CENTERS FOR DISEASE CONTROL  
1600 Clifton Road  
Atlanta, Georgia 30333**

#### BATTELLE DISCLAIMER

This is a description of research performed for the United States Government by Battelle. Because of the uncertainties inherent in research work, Battelle assumes no responsibility or reliance upon the information contained herein, beyond any express obligations embodied in the governing written agreement between Battelle and the United States Government.

## ACKNOWLEDGEMENTS

Battelle wishes to express special thanks to the large number of public health professionals at CDC and in states and territories, who provided us with information for this study. It is their experience and their perception of the role of computers in public health assessment which forms the basis of this study, although the opinions expressed here are solely those of the authors.

We extend our appreciation to Barbara Panter-Connah of the Epidemiology Program Office who acted as technical monitor for this project. She was as helpful to us as she was reported to be by the state and territorial epidemiology staff interviewed in the course of the study. We are grateful for the support of Floy Cross of the Office of Program Planning and Evaluation who was the project officer and to Nancy Chalmers of the Epidemiology Program Office who acted as co-technical monitor. Finally, our thanks go to Steven Teutsch, Melinda Wharton and Kathy Rufo who were vigilant about keeping us updated on new surveillance initiatives at CDC and on events in the Surveillance Coordination Group. Dennis Raichart of Battelle provided us with valuable insights into technical aspects of database management and computer support of surveillance.

## EXECUTIVE SUMMARY

**TITLE.** An Evaluation of the National Electronic Telecommunications System for Surveillance (NETSS)

**CONTRACT NUMBER:** 200-88-0642

**SPONSOR:** Centers for Disease Control  
1600 Clifton Road  
Atlanta, Georgia 30333

**CONTRACTOR'S NAME**  
**AND ADDRESS:** Battelle Memorial Institute  
2 101 Wilson Boulevard, Suite 800  
Arlington, Virginia 2220 1

[Pages 1-2] **I. Statement of the Problem**

The purpose of this evaluation study was to assess the effectiveness of the National Electronic Telecommunications System for Surveillance (NETSS) in facilitating the movement of surveillance data between state health departments and the Centers for Disease Control. The study examined the fit of present and planned NETSS developments with trends in the computerization of the public health assessment function at CDC and in states and territories.

NETSS is a system of computerized record formats which is presently used to transmit National Notifiable Diseases Surveillance System (NNDSS) data between CDC and the offices of state epidemiologists. Users of NETSS employ a variety of software systems to compile data from the state's own data management system and transmit them to CDC using an electronic mail facility or direct transfer. Data are uploaded to the NNDSS database on the CDC mainframe which support production of tables and graphics in the weekly MMWR and in annual NNDSS summaries.

NETSS began in 1984 with the inception of the Epidemiologic Surveillance Project (ESP) in six states. The prime

objective of the ESP project was to assess the feasibility of replacing weekly telephone reporting of general morbidity data from states to CDC with transmission of case-level data over an interactive computer linkage. In October 1989, all 50 states, Washington, D.C., New York City and Puerto Rico were transmitting data to CDC over the system. In the same year, ESP was renamed the National Electronic Telecommunications System for Surveillance (NETSS) to mark its emergence as a national system for routine collection of surveillance data from states and territories.

While the evaluation reported here was the first one to be conducted of NETSS itself, EPO has commissioned two related evaluations in recent years. A 1986 evaluation of ESP by Preventive Medicine Associates in 1986 noted the diversity of the process of computerization in the states. An evaluation of systems of infectious disease surveillance maintained by CDC completed by Battelle in 1990 found that state epidemiology staff generally felt that NETSS worked well and had at least the potential to streamline surveillance and reduce reporting burden in their departments. However, problems have persisted in timely data flow between states and the CDC mainframe environment. In addition, new developments to expand the capacity of NETSS to collect program-specific data, to introduce remote entry of data by local health agencies, and to provide access to new means of data analysis and dissemination are emerging at CDC and in states and territories.

In early 1990, The CDC Surveillance Coordinating Group (SCG) appointed the Subcommittee on Electronic Systems for Public Health Surveillance (hereafter referred to as the SCG subcommittee) to consider methods for improving the compatibility of surveillance systems designed by CDC and to develop policy recommendations for improving CDC support of the public health

assessment function in the future. Recommendations were approved in March, 1991. Strategies for the technical implementation of these recommendations have **been developed by the SCG.**

Page 1]

## **II. Evaluative Objectives**

This evaluation sought to build on these previous studies. Its overall goal was to support CDC program planning for future development of NETSS and other electronic surveillance initiatives at CDC and in states. The evaluation had three objectives:

- To see how the NETSS and related EPO activities operate at the present time to meet current surveillance needs at CDC and in the States,
- To see how planned enhancements to NETSS and other EPO activities will meet present and future surveillance needs at CDC and in the States, and
- To identify State and CDC needs which are not met by current or planned activities.

Recommendations for future support of surveillance were based on the findings of investigations in the EPO, other programs at CDC and in six states and territories chosen for this project.

[Pages 1 O-I 1 ]

## **III. Methodology**

This was a case-study evaluation based on interviews with staff at CDC and in six state and territorial health departments. The study had three components: 1) establishing a baseline description of NETSS in interviews with EPO and IRMO staff who worked on its development, 2) investigating the present and future role of NETSS in meeting surveillance needs of other C/I/Os at CDC in interviews with CDC staff managing surveillance systems, and 3) examining the present and future role of NETSS in state and territorial health departments.

The primary source of data, both at CDC and in states, was interview data from public health staff who are users and operators of NETSS or who are otherwise involved in surveillance. Interview data were supported with written materials and direct observation of the operation of NETSS. A written protocol of interview questions was used in all state/territorial interviews to ensure that data from all sites were comparable and analyzable. Interviews, written documentation and observational data were compiled into reports of activities in programs and/or states. These reports were reviewed for accuracy and completeness by persons contributing data to their development. Analysis was performed with a comparative methodology in which conclusions were drawn on the basis of differences and similarities in data on key issues across staff from varying agencies, positions and states/territories.

#### **IV. Major Findings and Recommendations**

[Pages 2-91

Findings and recommendations were structured around six key topic areas which emerged in the development of this study and in the SCG subcommittee recommendations for CDC support of electronic surveillance. These are: **1)** standardization/customization of software, **2)** training and technical support, **3)** software development, **4)** telecommunications, **5)** data exchange and connectivity, and **6)** dissemination. In addition, we address developmental priorities for NETSS, and inventory the resources available in states and territories to accommodate planned enhancements of the system.

[Pages 65-661

Standardization and customization. EPO would like to move to a more standard version of NETSS to improve support and make system upgrades more timely. The strategy followed by EPO since the inception of NETSS has been to customize the installation of the system for each user in states and territories. Customized

development of NETSS requires customized updating and operational support, imposing a limitation on the expansion of the system to a larger number of locations. In addition, states must often wait a long time for on-site upgrades of Epi Info applications.

State and territorial staff interviewed for this project did not reject the idea of a standard implementation of NETSS. Persons interviewed generally understand the need for a standard format to guarantee comparability of data from different reporting jurisdictions and to permit better quality control. Some epidemiologists supported the idea of greater standardization if it would lead to better technical support.

[Pages 74-761

Recommendations concerning standardization and customization:

- CDC should proceed with development of a more standard NETSS package. Development should include documenting the system as it is presently implemented in states and territories, defining and communicating the standard to be adopted and implementing the standard incrementally with short-term benchmarks.

[Pages 66-681

Training and technical support. For computer software to be used effectively, it must be properly installed, staff must be trained to use it, and there must be prompt access to someone knowledgeable about the software if something goes wrong. There are two models of the relationship of support to the adoption and use of software. One is to build specialized applications which maximize the capacity of the software to perform specific functions, but which require intensive technical support. A second model is to write a software package that is simple enough for the user to install and operate without intensive technical support and training. EPO has followed the first strategy with NETSS and EPI Info. EPO staff visit states and territories to install and upgrade Epi Info and provide ongoing telephone and Carbon Copy support.

States and territories place heavy reliance on EPO for technical support of start-up, ongoing operation and staff training for NETSS and Epi Info. This situation is unlikely to change in the near future. State epidemiology offices do not always have access to staff with personal computer expertise. However, there is evidence that states and territories are developing their own expertise in personal computer software and in Epi Info.

[Pages 76-8 1 ]

Recommendations concerning training and technical support:

- Providing a disk with an installation program and no further assistance or explanation is not adequate support for the kinds of major revisions envisioned by EPO.
- EPO should be prepared to provide assistance in installing and using software sent directly from IRMO to states if no other support is available.
- EPO should continue its on-site installation of Epi Info and NETSS in states and territories for the immediate future but should focus on capacity-building in states as a long-range goal.
- Training and technical support should rely less on “turnkey” systems and seek to educate the states on the system they are using.
- CDC should create a computer development internship to provide long-term computer expertise to states with special needs.
- CDC training in Epi Info should focus on training trainers to provide support to local health departments in the states.
- CDC should promote the teaching and use of Epi Info software in Schools of Public Health and other forums.
- CDC should support development of a mechanism for exchange of NETSS and Epi Info support among states.

[Pages 68-69]

Software development. The development of Epi Info, originally developed to support epidemic investigations and later applied to the ESP project and NETSS, has had dramatic effect on the direction of computerization of infectious disease surveillance in states and territories. In several states and territories visited for this project, Epi Info has changed the way in which surveillance is done by improving the access of health department staff to their own data. States can now analyze more data and they can do it much faster than was once possible.

Epi Info seems to be establishing itself as a standard for the management and rapid analysis of surveillance data. Epi Info was used for routine management and analyses of notifiable disease information in all the health departments visited. Even the non-Epi Info states chosen for this study, use it for some purposes. It is especially significant that it is used in all observed cases of local data entry. An increase in this practice may push states further toward an Epi Info standard.

[Pages 81-83]

Recommendations concerning software development:

- New or enhanced software developments should not be introduced into states and territories until both the software itself and specific applications have been well tested.
- Resources should not be devoted to competing with commercially available software that may already be in use in the states.
- Modifications to Epi Info to accommodate new operating systems and environments should be considered on a case by case basis.
- A fully LAN-Compatible EPI Info should be given a high priority.
- Epi Info enhancements should serve data management needs or specific public health objectives.

[Page 691

Telecommunications. The feasibility of direct transmission of data to CDC, timely transmission summaries and rapid dissemination of the data depend on development of a two-way capability in WONDER by IRMO. **The SCG recommendations** call for creation of a telecommunications “gateway” between reporting sources and the CDC communications networks via the CDC mainframe. This gateway will be developed and supported by **IRMO** with assistance from a technical advisory group. This proposal is responsive to requests by the Council of State and Territorial Epidemiologist (**CSTE**) for a single source at CDC for all disease reporting. Because so much of planned NETSS activity depends on this capability, it is essential that EPO and **IRMO** work closely together as this development proceeds.

[Pages 83-841

Recommendations concerning telecommunications:

- CDC should designate a team including representation from EPO and **IRMO** to oversee development of a telecommunications gateway between the CDC mainframe and remote users of NETSS.
- The **EPO/IRMO** team should review and update the schedule and milestones for the telecommunications gateway at least twice a year.
- There should be a joint beta test of NETSS and WONDER as soon as possible and this should be repeated at strategic points in the development of both systems.
- EPO should actively assist **IRMO** in developing the “gateway” to the CDC mainframe computer by providing technical assistance to states in the use of PC WONDER.

[Pages 70-721

Data exchange/connectivity. A network of linked computers in which all public health professionals can potentially have access to all others would provide timely access to public health data in a way that realizes the maximum potential of computers and telecommunications to make the right data instantly available at

the location where it is needed. Such a network will become a possibility with development of the CDC telecommunications “gateway”. However, development of such a network implies linkages of many kinds of reporting units including CDC programs, multiple agencies in state and local health departments, hospitals, clinics, providers, and any other involved groups.

Distributed data entry from district and local health departments is a high priority in states and territories. From the perspective of states, distributed data entry not only reduces data entry burden by decreasing the number of cases entered by any one agency, but will improve data quality by moving error resolution closer to the source of information. There is a size effect in the move toward distributed data entry. Small states have no interest in it and plan to retain control of data entry at the state level.

[Pages 85-861

Recommendations concerning data exchange/connectivity:

- Caution should be exercised in enlisting other CDC programs to adopt NETSS for their transmission. Agreements between program offices and EPO should be clearly defined and enumerated.
- CDC initiatives to encourage distributed data entry of infectious disease data in local health departments should be continued.
- Inter-state and direct intrastate connectivity is a low priority for states and territories visited in this evaluation, and should not be a high priority for CDC.

[Pages 72-731

Dissemination. Timely dissemination of data is a high priority for CDC and for public health officials involved in disease control programs. Surveillance data are “data for action” as well as for documentation of health events. If they are not available in time for action, their usefulness is reduced. Access to national

surveillance data in a timely fashion has been a deficiency of many CDC surveillance systems for infectious disease. Computers have the potential to improve the performance of CDC in providing rapid turnaround of surveillance data.

The usefulness of infectious disease surveillance in states is improved by rapid collection and dissemination made possible by computer transmission of data because it improves the access of states to their own data. However, the advantages to states of rapid turnaround of national surveillance reports are less clear. National surveillance is neither timely enough to support the control of infectious disease at the state level nor is it needed for this purpose. The rapid turnaround of data may be less important than is better quality control and analysis capability.

[Page 861

Recommendations concerning dissemination:

- The most useful rapid turnaround summaries for states and territories would be line-listed reports of transmissions which would permit ongoing reconciliation of data sets.

[Pages 86-891

**Developmental priorities**

There are a host of upgrades and enhancements to NETSS in various stages of planning and implementation at EPO which need to be prioritized. We recommend the following priorities for immediate development.

- First priority: The ADATABASE/NATURAL re-write of the NETSS system at CDC should be given highest priority in terms of resources and in terms of a focus for planning.
- Second priority: Detailed documentation of the NETSS system as it is presently installed in the states and territories should be EPO's second priority .
- Third priority: The degree to which state installations can be standardized, from the use of common Epi Info programming to a standard reporting form, should be determined and a standard developed.

- Fourth Priority: The standard implementation of NETSS then should be introduced to states and territories.
- Fifth priority: EPO should do everything possible to promote an effective telecommunications “gateway” between states and CDC.
- Sixth priority: CDC should implement measures to provide adequate technical support to states seeking distributed data entry at the county level.
- Ongoing: Software development priorities should be established on the basis of compatibility with the standard NETSS implementation and should not compete for resources with activities needed to maintain surveillance.

[Page 731

### **Resource Needs in States and Territories**

States and territories have or will soon have the hardware necessary to support present and planned enhancements to NETSS. All of the state and territorial health departments visited have computers with at least a 286 processor. Three of them have 386 processors. Disk space is adequate to run the Epi Info software, although storage of year-to-date NETSS data may become a problem in states with a very large number of cases. The data transmission hardware and software needed for a direct link to the CDC mainframe is already present in state and is used to send NETSS transmission to DIALCOM.

## TABLE OF CONTENTS

ACKNOWLEDGEMENTS .....	iii
EXECUTIVE SUMMARY .....	iv
CHAPTER 1. INTRODUCTION .....	<b>1</b>
Introduction .....	1
Background .....	1
Issues to be Addressed in the Report .....	2
Description of Methodology .....	10
Organization of the Report .....	11
CHAPTER 2. DESCRIPTION OF EPO COMPUTER INITIATIVES .....	13
The History of NETSS .....	14
Description of the Present NETSS Operations .....	17
Planned Enhancements to NETSS. ....	22
The Role of IRMO .....	26
Summary of Important Principles about NETSS .....	27
CHAPTER 3. APPLICATIONS OF NETSS BY OTHER CDC PROGRAMS .....	29
CDC Programs Reviewed .....	29
Acute Infectious Disease Surveillance .....	29
Chronic Disease Surveillance .....	32
Laboratory-based Surveillance .....	34
Other Health Data Systems .....	36
Issues Emerging in Study of CDC Programs .....	38
CHAPTER 4. THE ROLE OF NETSS IN STATES AND TERRITORIES .....	45
Criteria for Inclusion of States and Territories .....	45
A Description of States and Territories Included in the Study .....	48
The Role of Computers in Surveillance in States and Territories .....	49
The Surveillance Process in States and Territories .....	50
The Present Role of NETSS and Epi info in States and Territories .....	53
Future Directions in Surveillance in States and Territories .....	58
An Inventory of State Resources to Support Surveillance .....	62
CHAPTER 5. CONCLUSIONS AND RECOMMENDATIONS .....	65
CONCLUSIONS .....	65
Standardization and Customization .....	65
Training and Technical Support .....	66
Software Development .....	68
Telecommunications .....	69
Data Exchange/Connectivity .....	70
Dissemination .....	72
Hardware and Software Resources in States and Territories .....	73

RECOMMENDATIONS .....	74
Standardization and Customization .....	74
Training and Technical Support .....	76
Software Development .....	81
Telecommunications .....	83
Data Exchange/Connectivity .....	85
Dissemination .....	86
Developmental Priorities .....	86
REFERENCES.. .....	90

#### LIST OF TABLES

Table 1. A Comparison of Issues Raised by SCG and In This Evaluation .....	3
Table 2. Present and Planned NETSS Related Activities .....	15
Table 3. Summary of Criteria Used by CDC to Select States for NETSS Evaluations .....	46
Table 4. Attractions and Suggested Improvements to EPI INFO Suggested By States .....	56
Table 5. Computer Equipment Available to NETSS and Disease Surveillance .....	63

## CHAPTER 1. INTRODUCTION

### Introduction

The purpose of this evaluation study was to assess the effectiveness of the National Electronic Telecommunications System for Surveillance (NETSS) in facilitating the movement of surveillance data between state health departments and the Centers for Disease Control (CDC). The evaluation had three objectives:

- To see how the NETSS and related activities of CDC's Epidemiology Program Office (EPO) operate at the present time to meet surveillance needs at CDC and in the states,
- To see how planned enhancements to NETSS will meet present and future surveillance needs at CDC and in the states, and
- To identify state and CDC needs which are not met by current or planned activities.

In this chapter, we present the background of the study and describe the issues which it addresses.

### Background

NETSS is a system of computerized record formats used to transmit National Notifiable Diseases Surveillance System (NNDSS) data between the offices of state epidemiologists and CDC. Users of NETSS compile data from the state's own data management system using whatever software is routinely used and transmit it to CDC using an electronic mail facility or direct transfer. Data are uploaded to the NETSS data base on the CDC computer. These NETSS data support production of tables and graphics in weekly MMWR tables and in the NNDSS annual summaries of notifiable diseases.

While the evaluation reported here was the first one to be conducted of NETSS itself, EPO has commissioned two related evaluations in recent years. In 1986, Preventive Medicine Associates (PMA) completed an evaluation of the ESP project (1). This evaluation noted the diversity of the process of computerization, even in the six pilot states, and urged a wide-ranging study conducted by a coalition of leading state, local and federal health officials over a period of three years.

No such large enterprise has yet been undertaken. However, in 1989, EPO contracted with Battelle for an evaluation of infectious disease surveillance maintained by

CDC with special emphasis on the role of infectious disease surveillance in state health departments, the reporting burden generated in states by surveillance activities and the identification of redundant parallel reporting (2). The Battelle evaluation was completed in May, 1990.

The 1989-1990 Battelle study found that NETSS worked well in most state epidemiology offices and had the potential to streamline surveillance and reduce reporting burden in states and territories. However, the procedure for centralized, single-source reporting to CDC had been slow to materialize. Delays in the implementation of needed hardware and software enhancements at CDC, especially of a functioning communications link with the CDC mainframe computer, has been a problem. States also had concerns about training and technical support of software products and rapid dissemination of surveillance data in a form which they could use to generate their own reports. In some states, projects were underway or in the planning stages to facilitate automated transmission of infectious disease data to the state health department from local health departments or from providers.

Early in 1990, the CDC Surveillance Coordinating Group (SCG) appointed the Subcommittee on Electronic Systems for Public Health Surveillance (hereafter referred to as the SCG subcommittee) to consider methods for improving the compatibility of surveillance systems designed by CDC and to develop policy recommendations for improving CDC support of the public health assessment function in the future. These recommendations were approved in March, 1991 (3). Strategies for the technical implementation of these recommendations have been developed by the SCG.

### **Issues to be Addressed** in the Report

The recommendations of the SCG committee so closely parallel the issues used to guide development and implementation of this study that we have chosen to consider them in interpreting our results. The issues addressed by this evaluation are broader than those developed by the SCG committee, but they cover the same points. This convergence implies that the problems to be addressed in the immediate development of surveillance activities have been correctly identified. The issues which have guided this evaluation are compared to the SCG committee recommendations in Table 1.

**TABLE 1.**

**A COMPARISON OF ISSUES RAISED BY THE SCG AND IN THIS EVALUATION**

<b>Final' Report</b>	<b>Technical<sup>2</sup>Workplan</b>	<b>SCG Committee<sup>3</sup></b>
<b>Standardization and customization</b>	<b>Standardization/flexibility Case definitions</b>	<b>Standard core variables</b>
<b>Training and technical support</b>	<b>Training/technical support</b>	<b>Technical support</b>
<b>Software development</b>		<b>Software development</b>
<b>Telecommunications</b>	<b>Centralized reporting</b>	<b>Telecommunications</b>
<b>Data/exchange connectivity</b>	<b>Data sharing Local participation</b>	<b>Data exchange between health agencies and states</b>
<b>Dissemination</b>	<b>Dissemination</b>	<b>Availability of national surveillance data</b>

<sup>1</sup> Issues addressed in this report.

<sup>2</sup> Issues identified in the final technical workplan for this evaluation. October 15, 1990.

<sup>3</sup> CDC Surveillance Coordinating Group Subcommittee Recommendations on Electronic Systems for Public Health Surveillance. January 18, 1991.

### Standardization and Customization

The SCG recommendations call for implementation of a core set of descriptive variables to be collected for all events under public health surveillance. Consistent use of standard definitions, coding conventions and categories for these variables by all reporting jurisdictions will facilitate comparability and improve data quality control. While data do not necessarily need to be collected in a uniform way, they should be converted to standard format before being analyzed at CDC. Core variables including age, gender, race, ethnicity, date of onset and geographic location of the event are already generated by NETSS. However, the issue of standardization in the development of NETSS must be taken beyond the question of the format of final records to tie into the question of standard data collection procedures.

There is a tradeoff between standardization and flexibility in the development and use of software. A standard product tends to be reliable because - once the initial development and testing phase is completed - procedures have been debugged during the course of repeated use. Support is simplified because new procedures and updates can be provided to all users at the same time. However, use of a standard product means that everyone must do the same things in about the same way. While any number of procedures can be built into a standard product, there is a cost in development and start-up. The cost-effectiveness of developing applications goes up with the number of potential users.

An original premise of the ESP project, and one which has continued with the implementation of NETSS, is the idea that preparation of data transmissions for CDC should be derived from the computerized disease surveillance system already used in the state. Since the inception of NETSS, EPO has customized the installation of the system in each state and territory. Screens and variable names have been taken from forms and computer data bases already operating. Data are converted to standard categories and coding schemes used by NETSS before they are transferred to CDC in a standard file format. Programs to generate the standard NETSS transmission from state-specific screens have been written for each Epi Info state by EPO programming staff. Customization of NETSS not only provides CDC with what it needs but improves surveillance in states and territories. This has been an important selling point the adoption of NETSS by the states.

EPO would like to move to a more standard version of NETSS to improve support and make it more timely. This would require states to use standard reporting forms, variable names and record format, although EPO plans include a "customizable" state-specific screen to accommodate data needs unique to individual reporting jurisdictions. This study investigated the degree to which this is feasible in states, the limitations that states perceive in implementing more standard surveillance procedures and the acceptability of standardization to state epidemiologists and their staff.

#### Trainina and technical support

Several kinds of support are needed for computer software to be used effectively. It must be installed on the user's hardware and the options designated by the user must be correctly engaged. Users must be trained in the capabilities of the software, the options available and troubleshooting. Finally, there must be prompt access to someone knowledgeable about the software if something goes wrong. These three elements - installation, training and ongoing technical support - are an integral part of the software package. They are not extras. This is true of commercial software packages and of software developed by CDC for public health surveillance.

There are two strategies for supporting the adoption and use of software. One is to build in specialized applications which maximize the capacity of the software to perform specific functions, but which require substantial technical support of installation and operation. This is the strategy which has been followed by EPO with NETSS and EPI Info. EPO staff visit states and territories to install and upgrade Epi Info. They have supported the implementation of NETSS in all states either in-person or on the telephone. EPO uses Carbon Copy software to modify programs in states and territories from CDC using a telephone link. Training courses in the use of NETSS with hands-on exposure to applications have been conducted for staff from states and territories. The drawback to this strategy is that it is costly for CDC in CDC staff time and travel, at least in the initial phase of system implementation. As state staff become more adept at using the software, the support function served by CDC may be streamlined and support costs reduced. Another problem with customization is that there may be delays in support and training which delay full utilization of software.

A second strategy is to prepare a software package that is simple enough so that users can install and operate it without intensive technical support and training. The operator's manuals and tutorials on disk are adequate to support most users. Technical support is available by telephone for problems and special needs, but most users seldom need this. This approach is often taken by commercial software developers, such as those who produce wordprocessing and spreadsheet packages. At CDC, this is the approach taken by the developers of the Public Health Information System (PHLIS) and the Rapid Surveillance Helper (RASH) system in the Division of Immunizations. RASH is presently being used in 54 immunization projects to report measles case data to the Division of Immunizations, CDC. The PHLIS pilot is operational in seven states and is expected to be in use in all 50 by the end of summer, 1991.

The SCG recommendation on training and support calls for "...the necessary ongoing technical support of...staff in state health departments that use electronic public health surveillance systems designed by CDC." This recommendation is consistent with either model of support, depending on what is considered "necessary". In its planning for NETSS, EPO plans to maintain its present practice of providing intensive on-site and telephone support. Installation of the standard NETSS package will be a one-week on-site activity to include consultation, program design, programming, testing and debugging.

As part of this evaluation, we looked at several aspects of training and support of NETSS and Epi Info. We were interested in how dependent states and territories are on the on-site installation of initial programs and upgrades by CDC personnel. We examined the ability of users to maintain ongoing operation of the system using their own resources and/or those provided by CDC on an ongoing basis. We also investigated the adequacy of training and documentation to support use of Epi Info by users with a limited background in computers.

#### Software development

The SCG report recommends that surveillance software designed by CDC be designed to have consistent user interfaces and that it should be able to organize data in such a way that it can be sent to the CDC mainframe using a single telecommunications gateway. This recommendation addresses a proliferation of surveillance software that has been a concern at CDC and in the states. For several years, the Council of State and

Territorial Epidemiologists (CSTE), which makes surveillance policy for states and territories, has requested that there be some control on the development and use at CDC of multiple incompatible software products.

CDC found itself in the business of developing software for public health surveillance soon after computers began to be used at the federal and state level for this purpose. Development of software appropriate for this function fell to CDC because potential users of these products form a group which is too small to comprise a market for commercial software. The initial target audience for NETSS and Epi Info consisted of 54 state and territorial health department offices responsible for notifiable disease reporting.

NETSS and Epi info arose **in an environment** in which there was a great deal of software activity. Some CDC programs and some states had already made investments in commercial software for their own use. In other cases, programs at CDC had developed their own software applications independently of one another. Many of these were incompatible with one another. This places a heavy burden on public health staff at the state and federal levels who must learn several kinds of software to maintain surveillance of multiple health events. Presently, there are 26 different electronic surveillance systems operating at CDC using 10 different kinds of software (3). Ten of these systems using 4 different software packages operate in state epidemiology offices.

The problem of the proliferation of software was addressed by Battelle in a previous evaluation (2). This evaluation has been directed specifically to NETSS and Epi Info. However, we looked at the role of NETSS in standardizing the processing of infectious disease data and assessed directions in which software development should go in the immediate future. We also examined the question of testing as part of software development and its impact on the acceptance of software.

### Telecommunications

CSTE has requested a single reporting source at CDC several times over the past decade. Specifically, CSTE wants a single telephone number for all communications with the CDC mainframe. **The SCG recommendations call** for creation of a such a telecommunications “gateway” between reporting sources and the CDC communications networks via the CDC mainframe. Development of such a telecommunications link is presently underway. The telecommunications situation at CDC is changing very rapidly at

the present time. The description of the telecommunications link presented here is current as of May 15, 1991, the date that data collection for this project was completed.

The communications gateway will be developed and supported by IRMO with assistance from a technical advisory group with representation from other CDC programs and from states and territories. The communications gateway will accommodate a two-way exchange of information between users and the mainframe. Protocol specifications for linking to the gateway will be built into all software developed by CDC and will be made available to commercial vendors and any other person proposing to exchange software with CDC. Protocol conversions will occur at the gateway and require no action from users. To avoid cost of the system to users, an 800 number for users is under consideration. There will be no cost to users for access to the system. The telecommunications gateway will replace present electronic and hard copy reporting by states and territories to multiple offices at CDC. NETSS data, presently sent through the DIALCOM system maintained by the Public Health Network, will be transmitted directly to the CDC mainframe using this facility.

Direct investigation of the issue of parallel reporting to multiple sources at CDC was beyond the scope of this evaluation of NETSS, although the previous Battelle evaluation of parallel surveillance identified this as a problem (2). Our work here was restricted to the adequacy of DIALCOM as a transfer mechanism for NETSS. However, two proposed improvements to NETSS - direct transmission of data and improved dissemination of results - are directly dependent on the operation of a telecommunications link to the CDC mainframe. We have investigated these issues in some depth.

#### Data exchange/connectivity

CDC envisions a network of linked computers in which all public health professionals can potentially have access to all others by phone lines. This would provide timely access to public health data in a way that realizes the maximum potential of computers and telecommunications to make the right data available in a timely manner at the location where it is needed. Development of such a network implies linkages of many kinds of reporting units including CDC programs, multiple agencies in state and local health departments, hospitals, clinics, providers, and other involved groups.

Such linkages are easily developed technically but may create political problems for state epidemiologists who have legal responsibility for disease control programs. Not unreasonably, they are concerned that infectious disease data of which they are unaware might leave the state. The SCG committee addresses this concern, recommending that all data originating in state or local health departments be sent to the state office responsible for disease control before or at the same time that they are sent to CDC.

Data exchange between local, state and federal agencies means that there are multiple data bases covering the same disease events. Since infectious disease surveillance data, and possibly other kinds as well, occur in provisional and final forms, these data sets can be expected to differ somewhat at any given time. As a practical matter, one of the data sets must serve as a "gold standard", i.e. a reference data set to be used to resolve inconsistencies in local, state and federal surveillance data bases. The gold standard must have two characteristics: it must be accessible to all users and it should be the most accurate data set available.

In this evaluation, we discuss the implications of computer telecommunications technology for the control of state data. We also investigated the status of intrastate communications linkages with other state agencies and with local health departments. The question of the gold standard raised by the committee was not specifically addressed in this project, but we discuss the implications of our findings for this issue in Chapter 5.

### Dissemination

Timely dissemination of data is a high priority for CDC and for public health officials involved in disease control programs. Surveillance data are "data for action" as well as for documentation of health events. If they are not available in time for public health action, their usefulness is vastly reduced. Timely access to national surveillance data has been a deficiency of many CDC surveillance systems for infectious disease. Much national surveillance data is solely of historical interest by the time it appears in print. In some cases, data have not appeared for several years after the close of the time period in which they were reported.

Computers have the potential to improve the performance of CDC in providing rapid turnaround of surveillance data. The SCG recommendation proposes a one-year goal

for the turn-around of surveillance data. In this study, we discuss the role of national surveillance data in public health action and the priority of rapid turnaround relative to other qualities of surveillance systems.

### **Description of Methodology**

This was a case-study evaluation based on interviews with staff at CDC and in six state and territorial health departments. There were three components to the study:

- Establishing a baseline description of NETSS in interviews with staff developing NETSS, including EPO staff involved in the development and implementation of NETSS. We also talked to IRMO staff and technical staff in other Centers, Institutes and Offices (C/I/Os) at CDC who are working on projects to add disease-specific surveillance to NETSS.
- Investigating the present and future role of NETSS in meeting surveillance needs at CDC. This was done in interviews with non-technical staff in C/I/O's who have already added surveillance functions of their office to NETSS or are considering this. We also interviewed CDC staff who are involved in surveillance but have not expressed an interest in using NETSS.
- Investigating the present and future role of NETSS in meeting surveillance needs in state and territorial health departments. Project staff visited six states and territories to interview health department staff and to observe the operation of computer systems used for surveillance in the states.

The primary source of data, both in states and at CDC was interviews with public health staff who are users and operators of the NETSS system, or who are otherwise involved in surveillance. Interviews were accompanied by direct observation of computer systems in operation whenever possible. In addition, we collected written materials that could clarify the current and planned operation of NETSS. Relevant materials included in-house memoranda and documentation of computer systems. In states, we solicited organizational charts, computer documentation and print-outs and samples of analyses produced by the state.

Up-front planning was a very important part of the design of interviews. Interviews were kept open to permit unsolicited information from interviewees. Such input is a major advantage of open-ended interview designs. However, this kind of interviewing

requires careful planning to insure that important issues were included and retained throughout the study. In addition, it was necessary to control data collection in states and territories to insure that the data from all sites were comparable and analyzable.

We developed two instruments to guide data collection. A set of questions emerging from the issue identification phase of the project was used to guide interviews. The first version of this list contained all of the issues to be addressed. However, the focus of questions evolved over the course of data collection as issues surfaced in earlier interviews were evaluated in subsequent ones. For state data collection, a standard protocol was constructed for use in all states.

Analysis of the data was done using a comparative method. Reports were prepared describing the positions of CDC staff and of state and territorial epidemiology staff on issues defined for the study. Reports were supplemented with material available from observations and written documents. These reports were circulated to persons interviewed for review and correction. Where disagreements arose, the interpretation of interviewees was accepted as the relevant one. Once reports were finalized, summary comparisons were made showing responses to key questions across staff varying in position, office or state. Generalizations were drawn from similarities and differences detected in these comparisons.

Persons interviewed were promised confidentiality and were told that specific comments would not be attributed to them. In preparation of this report, we have endeavored to honor this commitment. However, this evaluation was done in a small world in which most people know each other. True anonymity is difficult to achieve in this context. A list of all persons contacted for this study is presented in Appendix A. The list of questions considered in the study and the protocol for the state site visits are presented in Appendix B.

## **Organization of the Report**

This report is organized to present the three components of the study separately with a synthesis at the end. Following this introductory chapter, we present a baseline description of the present NETSS system and proposed enhancements in Chapter 2. Chapter 3 contains results of the CDC study of present and planned use of NETSS for collection of program - specific data. In Chapter 4, the perspective obtained from visits to

states and territories is presented. We have developed conclusions for each of the three components of the evaluation in the last section of Chapters 2, 3 and 4 respectively. Chapter 5 presents conclusions which can be drawn from the study as a whole, and makes recommendations for the strengthening of NETSS and of CDC support of surveillance.

## CHAPTER 2. DESCRIPTION OF EPO COMPUTER INITIATIVES

This chapter discusses the purpose of NETSS and describes its implementation by the Epidemiology Program Office (EPO) which operates it at CDC. We will describe the emergence of NETSS, its present operations and plans for future enhancements. Four aspects of NETSS operations are discussed:

- Implementation in states,
- Transmission of data from states to CDC,
- Data management at CDC, and
- Feedback and dissemination from CDC to states.

The information presented here comes from interviews with EPO staff and representatives of the CDC Information Resources Management Office (IRMO).

NETSS is broadly defined in this study to include not only remote transmission of surveillance data but other computer applications for surveillance in localities, in states and at CDC that EPO has or will soon have under development. Narrowly defined, NETSS consists of standard categories, coding schemes, record layout and a telecommunications protocol which is used to electronically transmit infectious disease surveillance data from the offices of state epidemiologists to CDC. NETSS can operate regardless of the software used to prepare and transmit the data from specific states. However, EPO also supports Epi Info software to facilitate the management and utilization of surveillance data. They are also working on several other software tools for use in surveillance such as Epi Map and Epi Workstation.

Much of NETSS is under revision at the present time and any "snapshot" of current operations will quickly become obsolete. Nevertheless, for the purpose of this evaluation, we need to distinguish what is already in place from what is not yet available. Following a brief section on the history of NETSS, the discussion is divided into description of NETSS as it currently operates and NETSS innovations which are in some stage of planning but are not yet operational.

A goal of this evaluation was to assess the effectiveness of present NETSS operations and the feasibility and acceptability of planned enhancements of NETSS. Following a brief summary of the history of NETSS, we present sections on present and planned NETSS operations in four areas: implementations of NETSS in the states,

transmission of data from states to CDC, data management at CDC and feedback/dissemination from CDC to states. There is a discussion of the role of IRMO in NETSS. The chapter concludes with a review of features of NETSS which are especially relevant to this evaluation. Current and planned NETSS operations are summarized in Table 2.

### The History of NETSS

There was a window of opportunity for the computerization of infectious disease surveillance in the early 1980s. CDC and some large states were already processing surveillance data on mainframe computers, but the use of microcomputers for this purpose was in its infancy. The development of software to facilitate data exchange among public health agencies could build on a process of innovation in microcomputer technology which was underway in the states. This is much simpler than grafting a new set of procedures on an established process. Moreover, development of a single method for managing surveillance data on microcomputers could help avert the independent adoption of multiple and potentially incompatible hardware and software products by various states and agencies.

The impetus for the development of NETSS came from the Computer Working Group (CWG) of the Council of State and Territorial Epidemiologists (CSTE) in 1982. It was the goal of the CWG to foster the development of computer tools for management and transmission of a core set of infectious disease data among public health agencies at the local, state and federal levels. Software was to be developed which would not require a high level of computer expertise to operate, would not be too costly for states and localities to use, and would not depend on large, expensive machines.

The effort to computerize the transfer of infectious disease surveillance began in 1984 with the launching of the Epidemiologic Surveillance Project (ESP) in six states. This was a joint effort of CSTE and EPO. The prime objective of the ESP project was to assess the feasibility of replacing weekly telephone reporting of general morbidity data from states to CDC ("MMWR" reporting) with transmission of case-level data over a computer link. ESP performed well in the pilot states and was extended to additional states over succeeding years of the project. In October 1989, all 50 states, Washington, D.C., New York City and Puerto Rico were transmitting data to CDC over the system. In

**TABLE 2.**

**PRESENT AND PLANNED NETSS-RELATED ACTIVITIES**

	<b>PRESENT</b>	<b>PLANNED</b>
<b>Implementation in states</b>	30-byte core record used to send MMWR data.	60-byte extended record for MMWR data.
	Program-specific data transmitted for a few diseases from a few states.	Program-specific data appended to core 60-byte record.
	Epi Info, Version 3 or Version 5, for data management and analysis in most states.	Upgrade all systems to utilize Epi Info, Version 5, Epi Workstation, LAN compatible Epi Info
	CDC/CSTE case definitions working well.	No change.
	Data entry at state level and in a few local and district health departments.	Widespread data entry and transmission from local health departments.
<b>Transmission to CDC</b>	Transmission through DIALCOM E-mail. Direct transmission to CDC communications computer from three states.	Direct transmission to CDC mainframe through PC WONDER.
<b>Data management at CDC</b>	Data downloaded to EPO PC then to CDC mainframe.	Data loaded to CDC mainframe through PC WONDER.
	Data storage on magnetic tape.	Data storage on disk.
	Access by sequential read of entire data tape.	Random access to disk-based files.
	MMWR updates produced by read of NETSS file.	Merge of MMWR and NETSS data.

**TABLE 2. (Continued)**

**PRESENT AND PLANNED NETSS-RELATED ACTIVITIES**

	<b>PRESENT</b>	<b>PLANNED</b>
<b>Feedback and dissemination to states</b>	Transmission summary received after a week lag.	Timely turnaround of transmission summary.
	Annual reconciliation of NNDSS data.	Ongoing reconciliation of NNDSS data using NETSS verification records.
	Most access to data by reading MMWR.	Provide access to NETSS data through WONDER.
	Maps sent to states.	Epi Map to permit states to generate their own maps.
<b>Training and technical support</b>	Custom installation of NETSS in all Epi Info states.	Move to a standard installation of NETSS in all states.
	On-site installation of Epi Info upgrade from Version 3 to Version 5.	Ongoing and to be completed within two years.
	Telephone support by EPO of Epi Info and NETSS.	No change.
	Carbon Copy support by EPO of Epi Info.	No change.
	Two NETSS workshops at CDC for surveillance coordinators.	NETSS workshops to be conducted every two years.
	NETSS workshop for state epidemiologists on analysis capabilities.	

the same year, ESP was renamed the National Electronic Telecommunications System for Surveillance (**NETSS**) to reflect its emergence as an established national system.

The original premise behind the development of ESP/NETSS was that states would electronically transmit line-listed notifiable disease data to CDC using any computerized disease surveillance system already used by the states. However, the computer capacity of State and territorial health departments was found to vary widely. As the program was extended beyond the original group of states, EPO began to provide software, training and technical support to state health department staff overseeing the transition from hard-copy to automated transmission of surveillance data.

EPO did not decide what software states should use to computerize. However, they facilitated the computerization of surveillance by developing Epi Info software to support data management within states and to automatically generate NETSS transmissions from the state data. Epi Info is tailored for specific uses in epidemiology and public health, and also offers many of the capabilities found in commercial software packages. It can create data entry forms, validate entered data, accommodate file management functions, import and export data files, and perform standard statistical and numerical analyses. Epi Info, Version 3, became available in January, 1988. Epi Info 5.01, the most current version of this software, became available in April, 1990.

#### Description of the Present NETSS Operations

Implementation in states. All states, the District of Columbia, New York City and Puerto Rico transmit some of their notifiable disease data to EPO using NETSS. All states use the **30-character** core record standard for NETSS to transmit weekly NNDSS data. There are variations in the specific diseases notifiable in individual states and included in the transmission. In several states, local data entry has been initiated by the state health department with technical assistance from CDC. Some data are entered in district, county, or local health departments in Idaho, Missouri, New Mexico, New York, Pennsylvania, Tennessee and Puerto Rico.

Epi Info screens to be used for data entry have been customized by EPO staff to meet the state's specifications. Epi Info is programmed to recode the data to CDC specifications and prepare the data set for transmission to CDC. In addition to the MMWR

data, several states use a modified record to transmit data needed by other programs. Spinal cord injuries from four states, influenza isolate data from 15 states, animal rabies species from 17 states, and salmonella data from Georgia are examples of such transmissions.

EPO has supported state staff **in the implementation and ongoing management of NETSS. EPO** staff have traveled to 36 sites (33 states, the District of Columbia, Puerto Rico and the Virgin Islands) to install Epi Info and to train staff in preparing NETSS transmissions. In states which do not use Epi Info, EPO staff conferred with computer staff in the state to create an interface between the software being used by the state and NETSS. This was usually done by telephone, although five non-Epi Info states have been visited.

EPO has prepared a manual of NETSS operation and has conducted two seminars at CDC for state surveillance coordinators responsible for preparation of **NETSS** data transmission. A third NETSS workshop for mid-level state epidemiologists was conducted in February, 1991 to discuss data analysis and to introduce proposed changes and enhancements to NETSS. EPO provides telephone support of states using Epi Info. Support includes use of Carbon Copy, a software product that permits EPO staff to access and modify state software directly from CDC. EPO has worked with several states to help them implement entry of data at local health departments using Epi Info screens.

Transmission of Data from States to CDC. Most states report line-listed data in the NETSS format through the electronic mail services of the Public Health Network on the DIALCOM communications network. Three states - Georgia, Louisiana and New Mexico - are using Telenet to send NETSS transmissions to the CDC Communications Computer maintained by IRMO.

DIALCOM provides temporary storage of data files until it is convenient for EPO to download them. This averts the "traffic management" problems of scheduling transmissions from 53 states, localities and territories during the first two days of the work week when MMWR data must be prepared. However, the present procedure requires two transfers of data - a download to personal computers at EPO and a second transfer of downloaded files to the CDC mainframe computer.

At the time the data for this evaluation were collected, it was believed that there was an 80-column line length limitation for transmission of records over DIALCOM. This limitation has been mentioned by CDC staff outside of EPO as a problem in using NETSS to support collection of their own surveillance data. It was recently discovered that the default line length for DIALCOM is 120 bytes and that this can be extended to 300 bytes using a **linesize** command.

Data management at CDC. At present, much of the capture of NETSS data at EPO is done manually. EPO staff retrieve NETSS files submitted by states from DIALCOM using a Crosstalk command file. Files submitted by the states are concatenated online and downloaded to a personal computer at EPO. The download uses Crosstalk's file capture function which retains in memory any data scrolled on the screen by the host computer when a listing of the data is requested. This is not a true communications protocol and there is no automatic checking for transmission errors. Accounting for all of the files submitted by the states and ensuring that incoming data files are added to the single file used as a download holding area is done by the CDC operator. A significant increase in the volume of data entering EPO would be very difficult to manage in this manner.

Data downloaded from DIALCOM include line-listed case data and aggregate data. Aggregate data presently are transmitted from states either as textual information or as formatted summary records. The new NETSS format includes a "count" variable for this purpose. This process is being implemented at the present time but is not operating in most states.

Before being uploaded to the CDC mainframe, the data from DIALCOM are run through a preliminary edit routine written in "C" programming language by EPO staff. This program separates the captured file from DIALCOM into a data file and a text file. Data lines beginning with a legal state **FIPS** code in the first two columns and containing only numeric data in the first 30 columns or records with 9 in the first two columns and numeric data in the first 30 columns are individual case data or formatted summary data. These records are uploaded to the CDC mainframe. Data lines with alpha characters anywhere in the first 30 columns or with control characters anywhere in the data line are separated into a text file which includes textual aggregate data as well as "bad" records with alpha or control characters which may have resulted from noise on the line. These

are printed and visually scanned. Textual aggregate data are entered onto mainframe files using the MMWR entry program used in the past for weekly aggregate reports received from states over the telephone. Cases which are found to contain erroneous alpha or control characters are referred back to the states with the Case ID number for resolution of errors.

All NETSS records are maintained in a SAS data set on a multi-volume magnetic tape. In addition to the NETSS data set, MMWR data are maintained on the mainframe as aggregate data corresponding to their published form. The MMWR data base is updated with aggregated NETSS data using the reporting week, or "MMWR week", as a key for updating. Finally, annual data which have been reconciled with states and are considered final are kept in the NNDSS data set. This data set supports preparation of the NNDSS Annual Summary.

Mainframe processing of NETSS data is handle-d with SAS programs which were developed on an ad hoc basis as NETSS grew. Programs to detect errors in codes or onset dates, to update the NETSS and the MMWR data bases, to output monthly maps for some diseases, and to produce transmission summaries are run on the CDC mainframe. States correct or update records already included in the NETSS database by transmitting replacement records. Corrections are made to the NETSS database by matching replacement records to original ones on the basis of Case ID and state, and overwriting the earlier record. A record is deleted by submitting a replacement record with zeroes in all fields except the Case ID and state. Records identified as corrections to earlier data are coded by MMWR week and the corrections to earlier MMWR data are made manually using a tabulation created by a SAS program which reads the weekly transaction file and the NETSS database.

Updating of NETSS records has been hindered by inadvertent reuse of case identification numbers and a lack of unanimity in states on the interpretation of the report date variable. The record update procedures depend on unique case ID numbers within each state. Records intended to be replacements or deletions are identified as such if an earlier record in the file with the same Case ID number and state can be located. Inappropriate updates occur when some large states begin renumbering cases at the beginning of each year or at the completion of the maximum cycle allowed by the field

length of the variable. This problem will be solved in the new NETSS system which allows for a longer case ID and links year with Case ID.

The definition of "MMWR week" from report date is also not uniformly used in all states. Three states - Wisconsin, Illinois and N. Dakota - transmit and report cases by week of onset rather than by week of report. This means that their MMWR week reflects events which occurred one to two weeks earlier than the MMWR week reported by other states. The extended NETSS format will contain a date type variable which will permit users to identify the date used to determine MMWR week so that this possible inconsistency may be considered in analyses.

The use of magnetic tapes to store and upload data has created several problems. SAS processing of the NETSS data set requires repeated sequential processing of the entire data set stored on magnetic tape. Access to the data becomes ever more cumbersome as the NETSS data accumulate. Presently there are over one million records in the data set, so that even analyses that are limited in scope are hindered by the growing size of the data set. Identifying a replacement record requires comparing the Case ID with all other IDs in the file. Execution times for analyses involving the entire data set approach two hours and most processing must now be done overnight. Updates are written back to the same tape volume containing the original record, endangering the integrity of records if a run fails for some reason. Repeated processing of tapes has resulted in physical damage to the tapes with loss of data.

Feedback and dissemination from CDC to states. A weekly transmission summary is generated from each state's transmission and mailed back to reporting states. The transmission summary tallies the number of records received and indicates how many records were updates, duplicates within the transmission or contained errors. The transmission summary also provides lists of those cases with invalid dates of onset or codes for disease, sex and race. Year-to-date summaries of line-listed data are currently reported back to the states in the weekly transmission summary. States are notified by telephone of records placed into the text file because of alpha or control characters, or they infer errors from the number of records in the transmission summary, or they become aware of errors by checking the total for their state published in the MMWR.

Many of the errors which occur in the data are left to be dealt with in the resolution of numbers between states and CDC which precedes production of NNDSS Annual Summaries at the completion of the reporting year. This leads some EPO Staff to suspect that the transmission summaries are not providing the information needed by states in a timely enough fashion for them to maintain ongoing error resolution and updating of the NETSS data base. This was confirmed in discussions with state staff. (See Chapter 4.) Providing transmission summaries immediately, rather than mailing them later in the week, would greatly improve their impact on data quality and their usefulness in data verification in the states. However, at the present time, producing transmission summaries requires reading the entire sequential file in overnight computer runs.

### **Planned Enhancements to NETSS**

Implementation in states. Epi Info, Version 5, has been produced and is being used in the states. Version 5 offers many enhancements to Version 3, including expanded analysis capabilities, true missing values, fixed decimal fields and relational file handling. These changes provide the flexibility needed to adapt programs to include data needed by other reporting systems considering using NETSS. EPO personnel will visit states to complete installation of the Epi Info upgrade over the next two years.

An extended core record, enlarged from the present **30-byte** record to 60 bytes, will be introduced in the near future and is expected to become standard over the next two years. The extended core record format is final and EPO staff are putting finishing touches on documentation. Specifications for software to receive the data at EPO have been developed and programming is expected to be completed by Summer, 1991. The extended core record will improve the usefulness of the data and the simplicity of data management at CDC and in states. A frequency field will permit reporting of aggregate data in the same format used for line-listed data, eliminating the special textual treatment of aggregate counts. The new format will also have a code for probable and confirmed cases relative to the case definitions recently adopted jointly by CDC and CSTE (4).

Use of verification records in the new **NETSS** format will provide states with the ability to maintain ongoing reconciliation of their data with the CDC data set.

Verification records will be included in the weekly NETSS transmission from states to CDC and will contain a year-to-date summary of cases for specific diseases or events as known to the state system. These will be compared automatically to CDC data and a reconciliation summary showing any discrepancies will be sent back to the state. Use of verification records will be optional for states. This procedure will not pinpoint cases in error. There will still be some detective work to be done in states. However, it will narrow the range of cases which must be searched to a single disease or event type.

Efforts to expand NETSS to include reporting of program-specific data will be ongoing in the future. Disease-specific information will be located in a program part of the NETSS records to follow the core data located in the first 60-bytes. Following transmission to CDC, the core record will be used to support MMWR analyses. Both the core record and the program-specific record will be sent to the appropriate CDC program for further processing. This procedure would improve access to data at CDC and reduce reporting burden in states. The use of NETSS records to accommodate reporting to CDC surveillance systems other than the MMWR **has been initiated** for hepatitis and bacterial meningitis.

Over the next several years, EPO would like to move toward greater standardization of software used by the states to process and transmit surveillance data to CDC over NETSS. EPO will continue to provide custom programming to states for special needs. To support a standard NETSS/Epi Info installation, states will be urged to modify their reporting forms and procedures to conform as much as possible to the standard format. This will have a cost in terms of disruption of routine and delays in reporting during the startup period. EPO recognizes that standard software must provide an advantage to the states to offset its startup costs.

Two innovations which might favor the adoption of a standard NETSS package would be development of a Local Area Network (LAN) compatible Epi Info and the development of an Epi Workstation. LAN-compatible software permits multiple **local** users from different program areas within the state health department to simultaneously work on different records in the same file. Neither Version 3 nor Version 5 of Epi Info is LAN-compatible. When an Epi Info file has been accessed by one user, it is closed to all others. EPO is currently programming a fully LAN-compatible version of Epi Info's data entry module.

The EPI Workstation - a "bag of tools" for state epidemiologists - would be another asset to states if a more standardized NETSS program were developed. The Epi Workstation would be a set of standard software modules, including Epi Info, Epi Map, graphics capability and desktop publishing capabilities. These could be accessed from a single directory and operate on PCs at the state and local levels. The Epi Workstation will be developed at CDC and is presently in the planning stages. EPO has acquired a 486 PC as the basis for a surveillance lab to contain a test version of the Epi Workstation. However, Epi Workstation will be designed to work on 286 and 386 computers.

EPO plans to conduct NETSS workshops for state surveillance coordinators on a routine basis every two years. This will provide first-time training in the event of staff turnover in the states and will permit regular re-training of staff in enhancements to the system.

Transmission of Data from States to CDC. A direct telecommunication link between states and CDC will eventually replace the present DIALCOM method of moving data to CDC. This is essential to support timely transmission summaries and dissemination of results to states from the CDC mainframe. An important step in this direction is a project presently underway in IRMO to develop two-way communication between PC's at the state-level and the CDC mainframe utilizing a direct telephone link. This capability will be built into the Wide Ranging Online Data for Epidemiologic Research (WONDER) information System. WONDER will support both transmission of NETSS data to CDC and production of timely transmission and/or reconciliation summaries to states. EPO will be working closely with IRMO to integrate NETSS into this system, with plans to pilot test the system by January, 1992. CDC is also interested in communication packages which permit PC to PC interfaces between states and local health departments, with other states and with CDC. Interconnection of PCs would permit NETSS to realize the advantages of a network in providing flexible linkages among many operating nodes. However, there are no specific initiatives at the present time to support states in development of intrastate connectivity.

Data management at CDC. Mainframe data processing by EPO is being completely redesigned. SAS programs written for NNDSS (MMWR) and NETSS data management tasks will be replaced with NATURAL programming that will integrate them into a single

system. Sequential files stored on magnetic tape are being converted to ADABAS disk files. ADABAS will provide direct access to files, removing the need for sequentially reading the entire database with each update. This eliminates overnight scheduling of practically all processing and brings on-line transmission summaries and other reports within reach. Since recent releases of SAS can access ADABAS files directly, SAS graphic and statistical capabilities will still be available.

Once the ADABAS/Natural rewrite is completed, no new data processing initiatives are planned. The focus will be on analytic programs and support of other CDC programs after the core NETSS module is completed. The development of the ADABAS data processing system for MMWR data has been slowed by a shortage of staff, especially of Natural/ADABAS programmers.

Feedback and dissemination from CDC to states. Initiatives to support feedback of data between CDC and states include development of timely transmission summaries and immediate access to MMWR data on the CDC mainframe through the WONDER system discussed below. In addition, graphic and mapping displays of data and year-to-date results are being developed.

A timely transmission summary is essential if EPO is to fulfill one of the promises it made to states early on in the development of NETSS i.e. that electronic monthly reporting would eliminate the need for the tedious reconciliation of counts needed to produce the NNDSS Annual Summary. If the NETSS were working up to its potential, annual reconciliation would be unnecessary. Rapid turnaround of transmission summaries means that reconciliation of CDC and state data could be ongoing with errors detected and rectified weekly as they occur. The one-week delay in the receipt of transmission summaries means that errors which might be easy to correct immediately must be followed up for more information. This has led to an accumulation of error. Reconciliation for the Annual Summary now takes almost as long as it did when comparison of hard-copy was the only means to resolving inconsistencies (about eight months).

## The Role of IRMO

IRMO has responsibility for the CDC mainframe. No discussion of computerization of any CDC function, most especially surveillance, can overlook the importance of IRMO to the success of NETSS. To date, IRMO has not been closely involved in the movement of data between CDC and states over NETSS. With the exception of direct transmissions of data from three states, the CDC terminus for NETSS transmissions has been a personal computer in EPO. However, EPO initiatives to improve transmission from and dissemination to states depend on reliable access to a communications "gateway", an effective link between the CDC mainframe and multiple user PCs. States must also be able to access adequate support of mainframe operations.

IRMO sees telecommunications as its responsibility and plans to support all CDC telecommunication needs through a single system requiring minimal installation, support, training and maintenance. IRMO's first approach to developing a universal communications facility was the Communications Computer. This is a network of microcomputers running under the UNIX operating system. This system was developed in lieu of commercially-available bulletin boards to provide file transfer to and from the CDC mainframe and to avoid creating multiple systems at CDC and partly to allow running mainframe programs while state personnel are on-line.

The present Communications Computer does not have this last capability and has presented difficulties in more basic telecommunication tasks. IRMO plans no further developments of the communications computer or any other UNIX-based system. The present Communication Computer will be retained because some states are using it to report to EPO and to transmit laboratory data to the CDC mainframe over the Public Health Laboratory Information System (PHLIS).

EPO and IRMO are collaborating on preparing current NETSS data in a format suitable for presentation in WONDER. Summary data have been placed in a series of table cells, each cell representing multiple records with matching characteristics for county, demographic categories and disease. Summary tables can be edited to exclude small cell frequencies and single cell information for race, ethnicity, sex and age is not available at the county level for privacy reasons. The effectiveness of WONDER's user interface and its usefulness to untutored users in the states will be determined as its use becomes more widespread with the incorporation of NETSS data.

## **Summary of Important Principles about NETSS**

The goal of NETSS is to improve the usefulness of the surveillance data by making it more accessible to public health staff who use it to make decisions. Several characteristics of the system have been emphasized in the development of NETSS and its promotion to agencies at CDC and at the state and local levels. Some of these affect the potential of the system to be extended beyond its present scope. We summarize these points in this section.

The implementation of NETSS in states and territories has been customized for each user.

This means that screens and variable names have been taken from existing data management practices in states rather than imposing a standard format from outside. The program to generate the standard NETSS transmission from state-specific screens has been written for each state.

User-friendliness has been an emphasis in the development of NETSS. NETSS has been designed as a “turnkey” system to be used without difficulty by people with little or no previous computer experience. Epi Info in particular has been designed and distributed with this in mind. In Epi Info states, NETSS transmissions are generated automatically from an Epi Info menu option. This means that the operations needed to generate the NETSS transmission are not visible to the operator in the state. Epi Info is also supplied with pull-down menus and a well-written manual and requires no special syntax. Epi Workstation is a proposed development which will enhance the usefulness of computer management of surveillance to users without significant computer training.

All training and technical support of NETSS has been provided by EPO. Up to the present time, EPO has been responsible for all activities related to the development and implementation of NETSS. This includes development of programs, installation of programs in states and territories, training in the use of NETSS and Epi Info, and ongoing technical support of these developments. EPO maintains separate technical support of customized installations of NETSS in 36 jurisdictions that use Epi Info.

Improved timeliness of surveillance is an important advantage of NETSS, although reconciliation of annual NNDSS data remains slow. The usefulness of surveillance, especially infectious disease surveillance, is improved by rapid collection and dissemination made possible by computer transmission of data. Not only is the initial assembly of data faster, but updating can be performed more rapidly by transmission of corrections. This improves the quality of guidance provided to public health officials by the data. Timeliness will be further improved by faster error resolution with timely transmission summaries and by the addition of local data entry capability. Annual reconciliation of data also will become more timely when rapid transmission summaries make ongoing weekly reconciliation of data practical for states.

Enhanced analysis capability is a feature of Epi Info which has been very popular with those who use it. Rapid production of routine and ad hoc analyses has immensely improved access of users to their data and permitted them to use it in ways that were previously difficult or impossible. Epi Info supports analyses routinely done by epidemiologists and permits easy programming of customized analysis by people without programming experience. Addition of the Epi-Map programs and better graphics capability will further improve the analysis capability of Epi Info.

Future NETSS developments are dependent on concomitant developments in IRMO. Direct transmission of data to CDC, development of timely transmission summaries and dissemination of the data over WONDER depend on the development of an effective communications gateway with the CDC communications computer. Both the feasibility of these innovations and their timing is related to progress made by IRMO in developing an implementation of WONDER which can both receive and disseminate data - the single "gateway" to the CDC computer. This may take a long time to become fully operational. While IRMO proposes to pilot test the system, presently scheduled for early 1992, the SGC Subcommittee on Electronic Systems suggests that notifiable disease data will not be linked to this gateway until December 31, 1993.

### **CHAPTER 3. APPLICATIONS OF NETSS BY OTHER CDC PROGRAMS**

One goal of EPO in the development of NETSS has been to expand the system beyond collection of data for the MMWR to accommodate surveillance functions of CDC programs outside of EPO. This activity could increase access to data and reduce the burden of infectious disease surveillance at CDC and in the states. In this chapter, we describe the data needs of CDC programs which provided us with interview data for this project. In the final section, we discuss issues which may arise in the expansion of NETSS to incorporate data collection functions in other Centers, Institutes and Offices (C/I/Os) at CDC.

#### **CDC Programs Reviewed**

Some CDC programs already have begun to transmit data over NETSS. Several states use a modified record to transmit data for spinal cord injuries, influenza isolates, animal rabies species, and salmonella. Program-specific data for two diseases, viral hepatitis and bacterial meningitis, are being transmitted over NETSS from two states to the CDC communications computer. Other programs plan to develop NETSS capability. Still others are exploring NETSS as one of several options for collecting surveillance data. We tried to identify programs for this study with a variety of needs and in a number of different stages of buy-in to NETSS.

There are important differences in the purposes for which public health data are collected by various CDC programs. These may affect the suitability of NETSS for the data, the amount of modification which will be needed to implement reporting of data over NETSS, and the feasibility of making such a transition. Programs reviewed for this study are placed into four categories on the basis of public health uses of the data and the kinds of support needed for data collection. These categories are acute infectious disease, chronic disease, laboratory-based surveillance, and other systems for health data collection.

#### **Acute Infectious Disease Surveillance**

Infectious disease surveillance supports efforts to control the spread of disease and to prevent outbreaks. The unit of reporting is the case of a disease. To detect excess

morbidity indicative of a potential outbreak, it is necessary to ascertain all incident cases of disease occurring in the population of a defined geographic area. Sentinel surveillance restricts the geographic area to some sub-unit, but still aims for total case ascertainment within the chosen area. Excess morbidity is determined by comparison to historical patterns of morbidity as reflected in confirmed cases. Therefore, accurate updating of case records is necessary. However, one does not need to track individuals over long periods of time. This is the type of surveillance that NETSS was developed to accommodate. Timeliness is very important and there must be ability to update records for confirmed cases. Four systems for surveillance of acute infectious diseases outside of EPO were examined in this evaluation and are described below.

#### Viral hepatitis

The Viral Hepatitis Surveillance Program (VHSP) is operated by the Hepatitis Branch in the Center for Infectious Diseases. VHSP monitors hepatitis A, hepatitis B, Non A - Non B hepatitis, and unspecified hepatitis. Carriers and chronic cases should not be reported. Identifying and removing carriers is difficult and generally done at the state level. This is the most time-consuming step in hepatitis surveillance. The detailed data required to document risk factors for hepatitis and to screen out chronic cases are submitted on a hard-copy form which is key-entered at CDC. Hepatitis program staff would like to use NETSS to improve access to the data and timeliness with which they are analyzed and disseminated. Two states, New Mexico and Louisiana, presently send hepatitis data to CDC in machine-readable form using Epi Info and the Communications Computer.

#### Bacterial meningitis

Surveillance of bacterial meningitis is maintained by the Respiratory Diseases and Special Pathogens Branch at the Center for Infectious Diseases. There is both passive and active surveillance. The passive system collects hard-copy forms for confirmed cases from physicians at local hospitals. Active surveillance involves intensive tracking of cases. Records from the active surveillance system are sent from state epidemiologists to the project officer at CDC in an Epi Info file on diskette.

Surveillance of bacterial meningitis has taken on special importance because of a need to evaluate the effectiveness of immunizations against common forms of the disease. Active surveillance has shown underreporting through the passive system. Program personnel are hopeful that NETSS will encourage more complete reporting of cases through the passive surveillance system. Epi Info data entry screens have been developed for meningitis and are in use on a test basis in New Mexico and Louisiana.

#### Immunizable diseases

The Division of Immunizations (DI) is responsible for a disease control program for diphtheria, measles, mumps, pertussis, poliomyelitis, rubella and tetanus. Data are required not only to initiate outbreak control activities but to assess the adequacy of immunization programs and direct remedial action where vaccine programs appear to be failing. This means that the system requires both timeliness and detailed data on exposure and vaccination status. Laboratory confirmation of diagnosis is essential in early cases although epidemiologic linkage becomes adequate in an outbreak situation where early cases have been documented.

The Division of Immunizations (DI) collects data on measles through the Rapid Surveillance Helper (RASH), a dBase system in use in 54 reporting jurisdictions. Data are transmitted over telephone lines to a computer bulletin board maintained by the program. DI is committed to the RASH system for measles surveillance and is not willing to consider an alternative in spite of duplicate reporting of measles cases to DI and over NETSS. However, DI staff concede that RASH cannot easily be adapted to other immunizable diseases, such as pertussis. They suggest that NETSS may be an option for electronic transmission of surveillance data for these diseases.

#### Acute Sexually Transmitted Disease

Sexually transmitted disease (STD) includes a number of chronic and acute conditions which require different kinds of data and different modes of data management. STD cannot be treated as a single entity for the purpose of disease surveillance. There are important differences between acute and chronic STD and between high-prevalence and low-prevalence STD. Some STD, such as human papilloma virus (HPV) and herpes, are

chronic diseases and should be treated as such in surveillance. Others are so prevalent that the preparation of case-specific data is both impractical and unnecessary.

One possible use of NETSS may be to support sentinel surveillance of high prevalence STD. Line-listed data are needed at the local level to support disease control for syphilis and congenital syphilis. However, sentinel surveillance of high-prevalence diseases, such as uncomplicated gonorrhea and chlamydia, combined with aggregate reporting from all locations is adequate to trace the demographic distribution of the diseases and detect increased incidence in specific locations.

Currently, data on syphilis and gonorrhea are transmitted over NETSS as line-listed information from seven states. Other states and territories report the aggregate number of incident cases provided by the state STD program as a text message at the end of the NETSS transmission. STD program staff see NETSS as a possible means for transmitting line-listed data from states for syphilis, congenital syphilis, PPNG and possibly for the Gonococcal Isolate Surveillance Project (GISP) discussed below.

### **Chronic Disease Surveillance**

Surveillance of chronic diseases is conducted in order to identify risk factors for chronic diseases, guide prevention programs and document the course of the disease for purposes of health care program planning. The unit of reporting is the person with the disease. For data to be meaningful, it is necessary to follow affected individuals longitudinally over time. Since risk factors for chronic diseases such as cancer may have long periods of latency before diseases are manifested, registries of at-risk persons are needed to understand the relationship of risk factors to health outcomes. Timeliness is of less importance than the ability to connect multiple records for the same individual using relational capabilities. Examples of chronic disease surveillance systems are cancer registries, environmental hazard registries and birth defects registries. In addition, there are some STD, including AIDS, which are best assessed as chronic diseases. We investigated two chronic disease reporting systems which are using or considering using NETSS to collect data. In addition, we briefly describe STD reporting for which a chronic disease model is more suitable than an infectious disease model.

### Spinal cord injuries

Surveillance of spinal cord injuries is maintained by the Division of Unintentional Injuries in the Center for Environmental Health and Injury Control (CEHIC). This surveillance is being conducted to test the feasibility of making injuries reportable conditions. Surveillance provides information about etiology, of injuries, an important element in the design of prevention strategies. Etiologies of non-fatal injuries are not available otherwise because hospital discharge data do not contain codes for cause of injury. Moreover, states already collect this information to support rehabilitation programs. All that is needed is to get it to CDC. In spite of the fact that there is long-term follow-up in states, CDC does not maintain longitudinal data on spinal cord injuries.

There is no standard reporting format. Injury reports are received from states on a quarterly basis in a variety of formats: hard-copy narratives, floppy disks, and telephone reports. Four states report injuries using NETSS and two more have expressed an interest in beginning to do so. States append 31 columns of data to the NETSS core record. CEHIC provides no support of states' use of NETSS, leaving it to states to put the data in the correct format. It is difficult to tell how this system will work since at the time of the interview for this study, no analyses had been performed due to a problem with the CDC NETSS tape. We understand that this problem has since been resolved.

### Lead Screening

Surveillance for lead toxicity in children through the collection of lead screening data is currently under study by CSTE. Surveillance will identify children at risk for lead toxicity and will locate risk in the population as indicated by elevated blood lead levels. This information will support policy decisions about measures to prevent exposure of children to sources of lead in the environment.

Surveillance for lead toxicity presents a task similar to that involved in establishing cases of chronic disease because it requires assembling multiple records on individuals. Cases are ascertained on the basis of programs to screen children believed to be at risk. Lead toxicity results from an accumulating exposure to environmental sources of lead. This means that a case may be detected after multiple screening tests for elevated blood lead levels, and that a child at high risk for lead toxicity may have multiple positive screenings. For surveillance data to be useful, it is necessary to know the number

of children at risk rather than the number of positive screenings. In addition, timely alerts of toxicity are less important than the identification of individual cases.

Lead screening is a two-step process: an erythrocyte protoporphyrin screening followed by a blood lead evaluation for those who screen positive. Laboratory results are processed through state public health laboratories. Identifying individual cases of lead toxicity from multiple laboratory tests requires identifying individuals, assembling laboratory results and adding demographic data to the record. To be most useful in clarifying risk, the record should contain a history of screening results including prior negative ones.

CEHIC staff estimate that compiling records for individual children would require up to an entire staff person in states which have a large number of children screened. This implies a considerable reporting burden on states. For this reason, CDC is working very closely with CSTE in assessing the feasibility of this reporting. CEHIC staff are considering NETSS as a mechanism to transmit the assembled data on individuals from states to CDC. Data from laboratories would be transmitted to the state for compiling using the Public Health Laboratory Information System (PHLIS) which has been developed specifically to accommodate transfer of data from state public health laboratories.

#### Chronic STD

STD program staff feel that NETSS is not designed for surveillance of chronic conditions such as HPV, genital warts and genital herpes. These tend to recur in individuals and a single case may be treated more than once. Reported incidence of these diseases is often related to program activity. STD program staff are satisfied with the present quarterly aggregate reporting of these diseases.

#### **Laboratory-based Surveillance**

This type of surveillance provides data to guide public health practice which depends on knowledge of the biological properties of a disease organism. Antibiotic sensitivities, therapeutic recommendations and immunization recommendations are examples of such actions. The unit of reporting is the laboratory isolate or specimen. This often represents a human case of the disease but may be an animal or a food or water source. Timely analysis and dissemination of aggregate data is important. Relational

capabilities are less important, as laboratory isolates are single observations and do not usually require follow-up of individual records. Two laboratory-based systems were discussed with CDC program staff as part of this study. Enteric disease reporting and animal rabies surveillance are other examples of laboratory based reporting.

#### Influenza isolates

The influenza surveillance program is maintained by the Division of Viral Diseases (DVD), CID. There are three influenza morbidity reporting systems: the influenza isolate system which collects aggregate laboratory data, a weekly assessment of influenza activity by the state epidemiologist, and a sentinel surveillance system for flu-like illness from about 150 physicians. Influenza isolate data are presently reported over NETSS from about 15 of the 30 states in which influenza is a notifiable disease. Five or six states also send aggregate data on postcards completed at state laboratories, and five more send only the postcards. Surveillance of influenza is highly visible. It is crucial that DVD have regular and reliable access to high quality data. Reports of influenza without state laboratory confirmation are meaningless because so many illnesses present similar symptoms. Also, case data are not useful without an identified influenza isolate that specifies the strain of influenza being reported and marks a case as confirmed by a laboratory. Accuracy of influenza isolate identification and timeliness are both important.

The amount of data transmitted over NETSS in support of this program is small, requiring at most addition of two variables to the core record. Nonetheless, there have been problems with the quality of the data being transmitted over the NETSS system. Tabulations of data received through NETSS reveal a large proportion of records with missing or invalid isolates. For some states, the isolate is missing for all records. Further, comparison of laboratory post cards and NETSS records for the half dozen states submitting both show discrepancies even when the isolate has been provided. DVD plans to convert to direct reporting of influenza isolates from the state laboratories using PHLIS.

#### Gonococcal Isolate Surveillance Project (GISP)

GISP is a sentinel surveillance system operated by the Division of Sexually Transmitted Diseases and HIV Prevention, CPS. Data are collected on the first 25 male gonococcal isolates identified at 21 participating public STD clinics. The goal of **GISP** is to

monitor trends in antimicrobial resistance in gonorrhea and to link these to patient demographics. NETSS has been proposed as a data collection mechanism for GISP. Data for GISP are presently submitted to CDC on disk.

### **Other Health Data Systems**

There were three health monitoring systems investigated in this study which have special data needs and do not fit into an easily identified category. These are not systems designed to detect or estimate all incident cases of a disease in some defined population, the classic definition of surveillance. The unit of reporting varies, as do the priorities and computer support needs of these systems.

#### **The Behavioral Risk Factor Surveillance System (BRFSS)**

The goal of the BRFSS is to estimate the prevalence of selected risk factors in the population at large from a random sample of individuals not known to be at risk. The BRFSS is based on self-reported personal behaviors collected in monthly telephone interviews with a random sample of respondents in 45 states and the District of Columbia. A three-stage cluster sampling design with random digit dialing is used and the interview is conducted with the first adult contacted at residences in the sample. The instrument is updated annually to be responsive to contemporary health issues. In spite of its name, the BRFSS is not properly speaking a surveillance system, but rather a periodic survey.

The interest of BRFSS staff in NETSS stems from informal conversations with EPO personnel. Some states continue to submit paper forms or use key-entry contractors while 23 others are using CAS-CATTY, a computer assisted phone survey and telecommunications package. NETSS and Epi Info are under consideration only for those states not currently using the CAS-CATTY system.

Adapting the BRFSS reporting system to NETSS would pose several challenges. Data are received from states in a variety of formats that require substantial editing and re-formatting at CDC. The record generated from each interview consists of a core block of 140 bytes plus seven optional modules addressing special interests of various states. The total record is about 350 characters in length to accommodate the maximum number of modules. The need for seven modules should pose no problem to the

relational capabilities of Epi Info, Version 5, but continuing changes to record formats could become a serious programming burden.

#### Women, Infants and Children (WIC) Program Data

The Division of Nutrition in the Center for Chronic Disease Control and Public Health Promotion (CCDCPHP) assembles data on pregnancies and child development collected by the WIC program from eligible women and children from 43 states, Puerto Rico and several Indian reservations. WIC data are collected for purposes of assessing the effectiveness of the WIC program in improving the nutritional status of children. The data are used to check on the adequacy of the "safety net" for poor children, and to detect patterns of deviation from growth norms and hemoglobin/hematocrit standards. Unit of analysis is the client (child and/or mother) in a specific reporting period. The same individual may appear at different clinics at different time periods. WIC data are used to support corrective action for program offices rather than to impel any public health action.

At first glance, the WIC data appear to be a good candidate for NETSS. The WIC program uses fairly a standard format for all reporting states. Data are entered in machine readable format and would not require data entry to transmit. However, the WIC data lack a standard record identification field that could be used to match and update cases. There is also a very large number of records to be transmitted. The 1989 files contained 4.5 million records.

CDC program personnel collecting WIC data are interested in NETSS as a way to integrate their program with other CDC programs, to standardize processing and to disseminate reports to states more efficiently. Program representatives want to know more about NETSS and consider plans to adopt the system to be in a very early and tentative stage. The immediate potential of NETSS for these data is affected by the size of the data set. This will require significant mainframe support.

#### Medical Examiner/Coroner Information System Program

This program, operated by CEHIC, assembles data on circumstances and causes of deaths requiring inquest or investigation from medical examiners or coroners in nine states and eight counties. Data are purchased by CDC and are often received as uncoded textual data and narratives or in a variety of locally-defined coding schemes. The

data require extensive reformatting and editing at CDC. Analyses of these data support development of model reporting formats for disseminating medical examiner data with use of text, tables and graphics.

CEHIC staff responsible for this program doubt that NETSS can be adapted to their needs. Quick turnaround of data and rapid processing is not a high priority for this system. Organizing data by date of report is not relevant. Finally, data are not coded at the reporting source in a fixed format suitable for transmission over NETSS.

### **Issues Emerging in Study of CDC Programs**

NETSS has the potential to streamline the movement of public health data between the locus of data collection and the CDC mainframe by providing a **single mode** of data transfer. However, this study suggests that the adaptation of NETSS to fill program needs may not be straightforward. Adoption of NETSS transmission by other programs may require flexibility in record format and substantial support of program needs.

#### Expansion of NETSS to surveillance of conditions requiring accumulation of information about a single individual will require adjustment of the record update procedures in NETSS.

Chronic conditions and environmental toxicity are examples of conditions in which the status of the individual changes over time in significant ways. NETSS was designed to support the collection of data for single cases of acute infectious disease rather than accretion of information about a continuing health event. Updating of NETSS records to reflect changes in the status of the case lead to overwriting the existing record with a new one. For conditions which accumulate data, new records for the same individual and updates of existing records must be distinguished and handled appropriately. This can be done with a combination of codes in the RECTYPE (new code for accumulating record) and UPDATE fields in the new NETSS format. It will not be possible to leave the management of multiple records to the program using the data, since these fields determine the initial handling of incoming NETSS records. The programming to manage new record types must be completed before these types of reporting are put on NETSS.

Not all program data are compatible with NETSS in terms of format.

NETSS was designed for the collection of data in a fixed core format. It is assumed that program data to be transmitted over NETSS will also have a fixed format within each category of disease, although the specific variables and record length may vary from one disease to another. The expanded NETSS format permits transmission of continuation records, but still assumes a standard record structure.

Some program data may not arrive at CDC in a fixed format. They may require variable length fields, or the capacity to transmit varying number of observations in a single field, or the ability to transmit long text records. Either the NETSS format must be adjusted to accommodate variable length or variable format records, or the burden of coding data into a fixed format must be moved to some source at the originating end of the NETSS transmissions. This would be a problem with data such as that in the medical examiners system with its reliance on text data which are coded at CDC.

Some program data may not contain standard identification fields which can be used to maintain and update records transmitted over NETSS. For example, most of the 43 states participating in the WIC program send machine-readable, unedited records on a monthly basis. Records may be identified by full names, partial names or identifiers that have been scrambled, and the use of unique ID numbers is sporadic. No single field of ten or fewer characters is present in the WIC data to serve as a record identifier for deletions, corrections, etc. as expected by the NETSS system.

Use of NETSS to transmit data sets with a very large number of records may overwhelm NETSS operations.

This point arises out of our consideration of the potential of NETSS to provide a linkage between WIC clinics and the CDC mainframe. The WIC program's 1989 volume of 4.5 million records is an order of magnitude larger than the entire NETSS data set at the present time. An increase of this order calls into question the capabilities of phone line data transmissions and of computer hardware in the states. The WIC program would plan to pilot NETSS in a few states before going to full implementation. However, a pilot study cannot address the basic issue of whether the sheer volume of records will overwhelm the system. It is instructive to note that some of the interest in NETSS on the part of WIC program staff is due to their losing ready access to tape drives. EPO should consider that

their commitment is implicit in what may turn out to be an attempt to replace an existing mainframe system. This applies not only to WIC data, but also to any other source of health data which has very large number of records.

CDC programs adopting NETSS will need technical support and training during and after the transition.

NETSS was generally viewed as easy to use and may be widely adopted at CDC because there is no cost to doing so and because CDC staff perceive that it will be supported by the Epidemiology Program Office. More than one of the program staff interviewed mentioned this as an advantage. However, support may create a problem, depending on the expectations of program representatives and the ability of EPO to provide additional support. The degree of support expected or required from EPO may be unrealistic. A lack of understanding of a program's requirements on the part of EPO and of NETSS operations on the part of the program may lead program personnel to expect more assistance than is available and EPO staff members to underestimate the difficulty of adapting NETSS to program requirements.

Support and training will be needed in three areas: installation, state/local training and management of data after transmission. Adapting a program to NETSS may require an extensive array of custom-tailored Epi Info programming, script files for communications programs, and DOS batch files to match the existing capabilities of programs. If there is a change to existing modes of data collection, installation of software in states or localities may be required. Some management of the data after its transmission over NETSS will be essential, even if only to direct an incoming transmission to the proper location on the CDC mainframe.

Installation, and the training and technical support that follow, could become an open-ended commitment with new personnel and expanding program missions. Programs adopting NETSS will need to produce the tables and reports now available through existing surveillance reports. As new analyses made possible by NETSS and Epi Info come into use and as programs with divergent interests are accommodated, reports will be less standard and customized reports will be more in demand. Coordination and training for local and state personnel for key entry and analysis may be required in any changeover, although this training may be assumed by CDC program staff. Finally,

training and support requirements will expand as more people, some with little or no familiarity with NETSS or with disease surveillance in general, are brought on board. For example, the WIC program and the BRFSS relate to an entirely different constituency in the states than do infectious disease programs which deal with state epidemiologists.

Less than optimal support of NETSS can be costly to CDC programs which adopt this mode of data collection and find that, for some reason, they cannot use the data. There have already been support problems for two programs currently using the NETSS system for relatively modest data requirements. Representatives for influenza surveillance have had inadequate notice of planned activities and enhancements to NETSS that affect their program. Coding schemes and computer passwords for NETSS data have been changed without notifying them in advance. Analyses of spinal cord injuries were delayed because of processing problems. These problems can be and were remedied, but not before they had caused some, disruption of program activity. This may become a larger problem with growth in the number of programs covered by NETSS.

Some programs eniov support from CDC that would be shifted to states if NETSS were adopted.

This may be a selling point for the use of NETSS at CDC, but at a cost to the states. For example, the hepatitis and bacterial meningitis programs presently receive several thousand paper forms each year that are key-entered onto the mainframe by IRMO or by a contractor paid for by IRMO. This activity will be assumed by states if NETSS is used to transfer these data. CDC program staff believe that states are amenable to assuming the responsibility for the key-entry of their data since this will improve their access to the data and will eventually eliminate the need for maintaining paper systems. Our observations in states supports this contention. In addition, many states already **key-**enter these data in the states. However, it may be perceived as an increased reporting -burden by some states,

CDC programs have investments that they may not wish to abandon in order to adopt NETSS.

NETSS is not being offered to other CDC programs in an empty market. Programs have already made investments in computerization which are working for them adequately, if not perfectly. A clear example of this is the use of RASH for measles surveillance. The Division of immunizations has invested heavily in RASH and an in-house bulletin board for the reporting of measles. They have no interest in adopting NETSS in place of a working system. This does not mean that DI has no interest whatsoever in NETSS. They are open to the possibility of using it for other immunizable diseases such as rubella and pertussis.

Other programs have already made software investments. The lead screening program plans to use the PHLIS system to transmit lead screening data from the state public health laboratories to a central state office where they will be compiled into records for individual children. They have already invested in development of custom software to register and assemble child-based records from laboratory reports. The BRFSS has invested heavily in CAS-CATTY, a computer assisted phone survey and telecommunications package. The WIC program has developed an in-house software package called PedNSS which states use to process data before sending it to CDC. Adoption of NETSS and/or Epi Info by these programs may imply a commitment by EPO to make the NETSS application meet the specifications of software packages already in use.

The need to protect confidential data and issues of data ownership may complicate the use of a common NETSS gateway by some programs.

Generally, stripping the data of identifying information such as name and address is sufficient protection of confidentiality. However, improved dissemination of data through WONDER can compromise confidentiality. The technology exists to protect access to data but some kinds of programs have special issues with confidentiality. Spinal cord injury surveillance is an example of a system in which confidentiality is a special concern. Since these injuries are often due to traumas arising from accidents, the data may pertain to cases under litigation which could be impacted by breach of confidentiality. In this case, tabulations revealing the very small cell frequencies for these injuries could permit the identity of cases to be inferred.

A related issue is that of data ownership. Most surveillance reporting to CDC is done on a voluntary basis. As program offices with different mandates and missions look to NETSS as a conduit for the transmission of data to CDC, uncertainty may arise over attribution, authorship, responsibility for data quality and authorization for the release of data. This was viewed by the lead screening program as a concern, but not an insurmountable problem. It is often viewed as a problem in states.

The rapid turnaround of data may be less important to some programs than better quality control and analysis capability.

Rapid collection, analysis and dissemination of surveillance data is a selling point for NETSS which has often been emphasized by EPO. This evolved from the MMWR's "early alert" function, which in turn oriented record processing to the reporting week. This may be less important to some programs at CDC than other assets of the system. Rapid turnaround is of little interest to programs which do not see a sentinel role for surveillance data. It is a low priority for lead surveillance and inappropriate to the Medical Examiner data because inquests or other investigations can hold up reports for months. Other program representatives are uncomfortable with orienting tables, reports and processing to the week of report rather than the week of onset. This may turn out to be a problem with other programs which do not collect data on a weekly cycle. The new NETSS format has a date type variable that will permit reporting by week of onset. However, this should be made clear in discussions of NETSS with CDC program staff.

Some programs value rapid turnaround as a data quality strategy and not as a priority in itself. It is of value because it facilitates more timely attention to data quality control and complete reporting. For example, while timeliness is important to influenza surveillance, quick turn-around of data does not serve the influenza program if it interferes with their need to accurately identify cases. The priority of quality control was mentioned for the bacterial meningitis program and for the Behavioral Risk Factor Surveillance System.

For other programs, the potential advantages of NETSS are unrelated to rapid turnaround. The hepatitis program looks to NETSS to facilitate their entire process of surveillance, contribute to error reduction and provide additional analyses, reports and tables. The analysis capabilities of Epi Info to simplify routine and ad hoc analyses is an

advantage that the hepatitis program will receive as a result of their participation in .NETSS. The BRFSS staff also mentioned the capabilities of Epi Info as an asset. Staff managing WIC data see NETSS as a way to more closely integrate their system with other CDC programs, to standardize processing and to provide data and reports to states more efficiently.

## CHAPTER 4. THE ROLE OF NETSS IN STATES AND TERRITORIES

This chapter will present the results of interviews conducted with state epidemiologists and their staff as part of this project. The goals of the state study were:

- To see how NETSS and Epi Info are being used to support surveillance at the present time,
- To inventory present resources available to support computerized surveillance at the state level, and
- To determine future directions and needs for computerized surveillance in states and territories.

Following a discussion of the criteria for selecting states and territories to be included, we describe the surveillance process as it operates in most of these states and territories at the present time. We discuss the current role of NETSS and Epi Info and suggest some general trends which were observed in this study. In the final section of the chapter, we present an inventory of hardware, software and staff resources available in states and territories. Interviews were confidential and we have been careful not to attribute information to individuals. However, we have introduced findings from individual states and territories where needed to support specific points.

### Criteria for Inclusion of States and Territories

Data collection in states and territories involved on-site interviews with staff and, where possible, observation of health department operations affecting surveillance. Because of the intensive nature of the study conducted in states and territories, a small number of them needed to be chosen for this study. States and territories to be included were selected by EPO staff using a set of criteria chosen to ensure variation on several dimensions of surveillance likely to be relevant to the use of NETSS and to the impact of proposed enhancements to the system. Within the groups of states and territories fitting these criteria, five states and one territory were chosen on the basis of judgement by EPO staff who are familiar with the state programs. The states and territories chosen and their values on selection criteria are presented in Table 3. Criteria for the selection of states and territories were as follows:

TABLE 3.

SUMMARY OF CRITERIA USED BY CDC  
TO SELECT STATES FOR NETSS EVALUATION

	MN	MO	NM	NY	SD	PR
Hardware/software Micro, Epi Info Micro, <b>non-Epi</b> Info Mainframe/Mini	dBase	✓	✓	P ✓	✓	✓
Source of technical support CDC staff State staff	✓	✓	✓	✓ ✓	✓	✓
Locus of data entry Central Distributed	✓	✓	✓	✓ P		✓
Record format used Core only Extended format	✓	✓	✓	✓	✓	✓
Size of case load Small Medium Large	1127 ✓	3796 ✓	1244 ✓	6152 ✓	251 ✓	3790 ✓
Means of data transmission <b>DIALCOM</b> Direct	✓	✓	✓	✓	✓	✓
Year began ESP	1984			1984		
Year began NETSS	1985	1987	1986	1985	1988	1986
Year of EPO visit	None	1987	1985 1988 1990	1990	1988	1985 1989 1990

\* Sum of cases in the 1990 NNDSS Annual Summary for AIDS, aseptic meningitis, primary encephalitis, hepatitis A, hepatitis B, hepatitis NANB, unspecified hepatitis, legionellosis, leprosy, measles, mumps, meningococcal infection, pertussis, rubella, primary and secondary syphilis, toxic shock syndrome, tuberculosis, tularemia, typhoid fever, Rocky Mountain Spotted Fever, typhus and animal rabies.

Hardware and software presently used for surveillance. This criterion was used to ensure inclusion of both Epi Info and non-Epi Info states, states using a mainframe or minicomputer, and of states and territories using personal computers for surveillance. Non-Epi Info states were selected to include both a mainframe state and a state using personal computers.

Source of technical support for surveillance software. Technical support of software can be an important determinant of success or failure. In order to assess the role of technical support in the performance of NETSS, states and territories which depend on CDC support of software were compared to states which do not. All states and territories using Epi Info are supported by CDC (EPO). States using a mainframe or a PC with some other software must provide their own technical support of software.

Locus of data entry. Epi Info and NETSS can easily be expanded to cover distributed data entry. States and territories were chosen in which data entry occurs at remote locations such as a district office of the state health department or a local health department in addition to states in which data entry is done at the central state office. In practice, this distinction was not clear-cut. In two of the states studied, remote data entry was occurring in only one or two locations.

Record format used. A NETSS initiative soon to be implemented is the processing and electronic transfer of program-specific data over NETSS using the relational capabilities of Epi Info, Version 5 and an extended NETSS record. One state was included which is using this capability at the present time to send records in the extended NETSS format. The other states and territories included transfer only core data using the **30-character** core record.

Size of case load. This is an important determinant of demands made on surveillance software and the difficulty of routine surveillance operations. Large, medium and small states may have different needs and problems in managing, analyzing and using surveillance data. For this reason, states and territories were chosen to cover a range of case load size. The choice of states and territories was based on judgement of EPO staff who work with states and territories, rather than on explicit size criteria.

Means of data transmission. Proposed changes in data management at CDC and in remote access to the CDC mainframe will permit direct transmission of data using a modem and a communications protocol. One state presently using direct transmission was chosen to permit evaluation of this dimension. The other states and territories chosen presently use DIALCOM to transfer NETSS data.

#### **A Description of States and Territories Included in the Study**

In this section, we discuss the positions of states and territories relative to selection criteria and present brief descriptions of surveillance operations in the states and territories which were visited. State descriptions are presented in rank order from the largest to the smallest number of cases.

New York. New York has a mainframe-based system with a large case load. It was one of the initial six states to participate in the Epidemic Surveillance Project (**ESP**) and began transmitting data to CDC in 1984. At present, key entry and most data processing is done with the state's mainframe computer system. New York has expressed an interest in converting to Epi Info on personal computers, but the transition will be a big job in this large state. Two counties have adopted Epi Info for key entry of cases at the local level and send their data to the Bureau of Communicable Diseases on floppy disks. New York City also sends copies of their data to the state health department, but also transmits MMWR data directly to CDC. Over the next several years, New York hopes to have all counties key enter their own data. They look to CDC for support in implementing a distributed data entry system using Epi Info.

Missouri. Missouri has a microcomputer-based system using Epi Info software. Data are entered on Epi Info screens at district health departments and transmitted to the central state office using PROCOMM and a modem. State staff compile the data into NETSS transmissions and send them to CDC over DIALCOM. Local health departments in St. Louis and Kansas City also transmit data to the district office where it is integrated into the district file prior to transmission to the state. Missouri began transmitting data over NETSS in January, 1988.

Puerto Rico. Puerto Rico uses a microcomputer-based Epi Info system. Data entry occurs at the level of autonomous regions and is delivered to the central health department office on disk weekly. Remote transmission of data from regions to the center over modems will be initiated in the near future. NETSS was introduced to Puerto Rico in 1989 and all data began to be sent to CDC via NETSS on January 1, 1990.

New Mexico. New Mexico has a microcomputer-based system using Epi Info software. Data for most of the state are entered at the state office in Santa Fe, although data entry at the district office in Albuquerque has been occurring since June, 1988. Several proposed NETSS enhancements were introduced to New Mexico in July 1990, including use of relational data screens for enteric diseases, hepatitis and meningitis and direct communication to CDC in place of DIALCOM. There have been some problems with accessing Albuquerque data using the new system.

Minnesota. Minnesota has a microcomputer-based system which uses dBASE rather than Epi Info and does not receive software support from CDC Surveillance processing and the NETSS transmission are done with a **Foxbase** system developed in-house. Data are entered at the state level and transmitted to CDC over DIALCOM. Minnesota was one of the initial ESP states, and has been transmitting data over ESP/NETSS since May, 1984.

South Dakota. South Dakota has a microcomputer-based system using Epi Info to manage and analyze surveillance data. South Dakota has used Epi Info, Version 3 since January, 1989 to prepare and transmit data through DIALCOM. They hope to adopt Epi Info, Version 5 and the expanded record format in the near future. South Dakota was chosen to be representative of states with a small case load.

### **The Role of Computers in Surveillance in States and Territories**

State epidemiology staff interviewed for this project were enthusiastic about the potential of computers to improve the quality of public health surveillance and to facilitate the use of surveillance data in public health action. The most frequently mentioned advantage of computers is the improved access to data provided by computer data management. This permits expansion of surveillance to other conditions such as injuries, environmental illness and chronic diseases. Decentralization of data entry is also

seen as a means of improving both data quality and access at both the state and local level. In addition, computer processing of data can ameliorate staff shortages occurring because of state hiring and budgetary processes.

However, computers can create their own problems. We were cautioned that computers are not a panacea for problems with surveillance itself. The usefulness of surveillance data are limited by the quality of ascertainment, investigation, follow-up and error resolution with or without computers. The effectiveness of decentralization of reporting is limited by training, support and hardware at local level. Finally, start-up and transitions can be disruptive in both large and medium states and territories. Changes to surveillance procedures must be made without interruption of this essential public health function. It is not possible to stop the system while software and hardware are installed, modified or debugged.

### **The Surveillance Process in States and Territories**

With some minor variations, the process of collecting, processing, analyzing and disseminating surveillance data is similar in all states and territories visited. We describe the surveillance process in general terms in this section. Not all states and territories visited perform surveillance in exactly this way. Important exceptions are noted in the text.

Data Collection. Data are collected by local public health agencies from providers and others who are legally responsible for notifiable disease reporting. The data are collected on standard morbidity reporting forms or cards from which they are entered at the state, district or local level. Before the data are reported to CDC, a determination is made as to whether the reported incident meets the CDC/CSTE case definition. This determination can be made by epidemiologists at the state level or by public health nurses at the local level. All states and territories visited use CSTE case definitions, and report that they work well. Cases are reported provisionally and may or may not be updated later if they fail to meet the case definition. New York does not update provisional data, but provides CDC with a quarterly report of final data consisting only of confirmed cases.

Data Quality Control. Data quality checks for invalid codes or duplicate records are performed before data are sent to CDC. States and territories reported that few errors are

detected afterward. Epi Info states and territories use the check features of Epi Info during data entry. In all states and territories, sorted line listings of cases are used to check for duplicate reports. Responsibility for data quality control usually rests at the state level even in large states, with local responsibility for investigation and follow-up.

Data Transmission. Reports are sent electronically over NETSS for general morbidity data, with aggregate counts for STD, TB and formatted summary records for influenza isolates. Actual transmission takes from one minute to thirty minutes. States and territories using DIALCOM have experienced no problems with it. New Mexico, the only state with direct transmission to CDC, also reported no problems with this mechanism. New Mexico maintains its DIALCOM account for other purposes and has no intention of discontinuing it.

New Mexico enters program-specific data for hepatitis, meningitis and enterics over NETSS using the relational capabilities of Epi Info 5 and transmits program specific data for hepatitis and meningitis. All states visited, including New Mexico, send **program-specific** data to CDC on hard-copy forms. Puerto Rico sends only core data to CDC over NETSS and does not submit hard-copy forms. Two health departments have some program-specific data, notably hepatitis and meningitis, computerized and ready to send over NETSS if the programming were available. RASH is used to send measles data from four of the six health departments visited.

Annual Reconciliation. Reporting to CDC means that at least two data sets exist for each state, one at CDC and one in the state. There are two common sources of inconsistency between these data sets: errors in transmission and differences in case records due to failure of record updates or deletions. Transmission summaries mailed from CDC are compared to the state's own transmission records for anything which appears odd. Transmission errors are rare and are easily resolved over the telephone with CDC.

Transmission summaries contain the following: **1)** summary totals for the current weeks transmission, **2)** a listing of cases with invalid codes for disease, sex, race or onset date, **3)** a listing of cases with duplicate case ID numbers, **4)** a year-to-date count by disease code. Records with non-numeric characters and records which are the wrong length are stripped out of the transmission and printed out on the transmission summary. These are not added to the NETSS data base. Records with invalid sex, race or onset date

are listed so that they can be corrected, but are also added to the data base with unknown values in place of the errors.

The usefulness of the CDC transmission summaries for ongoing reconciliation of state and CDC data is limited because there is no way to identify the case or cases responsible for inconsistencies in aggregate counts. Difficulty in identifying erroneous records is a barrier to reconciliation rather than the fact that transmission summaries are not available immediately. The one-week delay in getting transmission summaries was mentioned as a problem in only one state.

All states and territories reconcile their data with that resident on the CDC database to support preparation of the NNDSS Annual Report. Annual reconciliation is a very important data quality control measure for the states and territories, but it is a big job in all states and territories which attempt to complete it. Some states and territories try to reconcile their data on a shorter cycle than the annual one required by the NNDSS. However, their capacity to do this depends on availability of a line-listing of information residing on the CDC data base.

New York State does not reconcile cases because of the large volume of cases which are reported. They maintain two data sets: provisional data sent in weekly over NETSS and confirmed cases sent to CDC on disk every three to six months. More frequent reconciliation is not feasible in New York, and more timely transmission summaries would not help. In all medium and small states and territories visited, health department staff would like weekly updates and corrections to avoid the massive annual reconciliation. Annual close-out and reconciliation has been a problem for states and territories which have recently adopted NETSS because procedures for completing the year-end transition were not clear to state staff. Annual summary reconciliation was discussed at both CDC workshops for reporters. However, this was not translated into easier reconciliation of cases in individual state systems.

Data Analysis. Monthly or bimonthly breakdowns of cases by demographics and geographic parameters are produced by all states and territories. These are normally circulated to the health community in the state in a newsletter format. In addition, annual reports are produced with tabulations, graphics and maps of case distribution across the state. Ad hoc analyses are also performed several times per week on request from other agencies or the public.

Epi info programs prepared by CDC or by state staff themselves are used to produce most analyses in the five states and territories visited with PC-based systems. Even Minnesota uses Epi info for ad hoc tabulations and lists. SAS and PC-SAS are also popular. Production graphics are done with commercial graphic packages rather than with Epi Info graphics. This is easy to do and produces higher quality camera-ready graphics. Three states were enthusiastic about Epi Map, especially if it gives states control of the analysis. For three others, it is not a high priority. Missouri was examining the possibility of buying or producing their own mapping program until they saw the presentation of Epi Map at the February NETSS workshop.

Data Dissemination from CDC. Data transmitted to CDC over NETSS are disseminated in the MMWR and in phone calls to program staff at CDC. In addition, EPO produces monthly maps of disease in contiguous states for six diseases. Annual data are more useful to state epidemiology staff than are weekly provisional data which are subject to fluctuations and effects of batch reporting. Annual data are not timely enough to support public health action, but are useful in discerning long-term trends **and** supporting comparisons such as those in MMWR Figure 1. There were some misgivings about the maps produced by EPO because of uncertainty about cut-points and denominators. Also the maps are not timely enough to support public health action by the state.

### **The Present Role of NETSS and Epi Info in States and Territories**

**There** are two kinds of conclusions to be drawn from the state study. These are related to the present status of the NETSS and Epi Info implementation in the states and territories visited and future directions for surveillance emerging from discussions with state epidemiologists and their staffs. In this section, we present perspectives on the current role of the system in state health departments.

NETSS has achieved the original goal of fitting into what states and territories already do to manage surveillance data. Narrowly defined as a mechanism for transmission of data for nationally notifiable infectious diseases, NETSS itself is transparent to state staff as long as it works well. Moreover, Epi Info has supported the tasks that states and territories must **do for themselves**. The question "what do you do just for CDC" was irrelevant to most state staff interviewed, especially in Epi Info states and territories. Staff

responsible for preparation of the NETSS transmission could not estimate the time spent specifically on this task. Production of transmissions to CDC is a low-profile part of routine surveillance operations and is not perceived as a reporting burden. Most states and territories would have developed or be developing some means to transmit computerized data even if NETSS were not there. It simply makes no sense to maintain both computerized and hard-copy data management systems over the long run. Minnesota illustrates the process of computerization in a state which has taken a course independent of Epi Info.

Changes in procedures for managing surveillance data are risky for states and territories because they cannot bring down the system to accommodate transitions. Communicable disease surveillance is an essential public health function in the states and territories which must be ongoing during any transition in the method for performing it. This is a problem in the change to NETSS and/or Epi Info from hard-copy data management. The transition to Epi Info and NETSS has been easiest in states and territories which had no prior computerization of surveillance. These states and territories found the transition to be smooth and the software easy to use. In states like New York, which have an operating mainframe system, the transition has not been completed and will be more difficult because an existing system must be disassembled and replaced.

States and territories are very vulnerable if they invest in a transition or an update which fails to perform immediately. Moreover, breakdowns can occur because of failure of equipment or staffing which occur frequently in state health departments. This can be seen in the recent experience of New Mexico in setting up data entry of cases from Albuquerque, the largest reporting jurisdiction in the state.

In 1988, EPO staff visited Albuquerque to convert the Albuquerque system from dBase to Epi Info, Version 3. Distributed data entry was introduced at this time and worked fairly well for a time. Problems developed as the Albuquerque file became too large to telecommunicate to the state health department in Santa Fe in a relatively short amount of time. To get around this, data were sub-setted for transmission to Santa Fe where they were to be merged into the state's master file. There were difficulties with the merge program which was corrected only after some delay. In addition, after the Epi Info version 5 installation in New Mexico in July, 1990, Albuquerque lost their computer support staff person and problems developed with the ID number variable in Epi Info, Version 5.

All of these problems left the state health department without access to data from Albuquerque, a large proportion of the cases in the state. These problems can be rectified and are being resolved at the present time. However, the disappointment and skepticism of staff who have encountered difficulties in accessing their own data may be harder to overcome.

Not all states need significant support from CDC for technical support of on-going operation of NETSS and Epi Info. Support of the adoption of NETSS and Epi Info enhancements was an issue in only two of the states visited here. In four of the states and territories visited, staff had no comment other than that they were satisfied with CDC support of NETSS operations. The states which rely the most on CDC were New York, a large state considering a switch from a mainframe system to Epi Info based reporting on PCs and New Mexico which is notable for its lack of computer support staff. The transition in New York will take time and will require staff resources which may be hard to come by in the state. New Mexico has no computer support staff to help them resolve immediate problems and must rely completely on long-distance support from CDC. Computer staff in some states and territories expressed reservations about sole reliance on the Epi Info manual to learn programming, and felt that there must be in-person training for state staff who have had limited computer experience. In Puerto Rico, training in Epi Info training of staff came mostly from courses at the School of Public Health at the University of Puerto Rico. Many of the people interviewed have attended CDC workshops and found them a very useful orientation to the system.

Epi Info has made a major contribution to the practice of public health in those states and territories that use it by improving their access to their own data. The development of Epi Info, originally developed to support epidemic investigations and later applied to the ESP project, has had dramatic effect on the direction of computerization of infectious disease surveillance in states and territories. In several states and territories visited for this project, Epi Info has changed the way in which surveillance is done -by improving the access of health department staff to their own data. States can now analyze more data and they can do it much faster than was once possible. We summarize state perceptions of the strengths and weaknesses of Epi Info in Table 4. In order to protect the confidentiality of responses while presenting the pattern of responses by state, we have replaced state names with randomly assigned numbers.

TABLE 4.

## ATTRactions AND SUGGESTED IMPROVEMENTS TO EPI INFO SUGGESTED BY STATES

	1	2	3	4	5
<b><u>Attraction:</u></b>					
Easy data entry			✓		✓
Easy access to data	✓	✓			
Easy to use for ad hoc data needs and analysis	✓	✓			
Portability (usefulness on laptops for field work)	✓				
<b><u>Suggested Improvements:</u></b>					
Graphs and other output are not presentation quality	✓	✓	✓	✓	✓
Graphs have too few options and styles			✓		
Column and row labelling inadequate	✓				
Range checking capability is inadequate for large number of city and town names			✓		
Does not support true double key-entry verification			✓		
No multi-user support				✓	
Unclear program diagnostics and error messages			✓		✓

Some of the improved analysis capability in states and territories is due to computerization itself. Almost any computer analysis is faster and more accurate than one based on manual sorts of hard-copy forms. However, by integrating data management and data analysis in a **single** application running on a personal computer, Epi Info has freed state **epidemiology** staff from dependence on central data processing units and the delays inherent in relying on mainframe support staff. This is especially important in the preparation of ad hoc analyses. For all practical purposes, ad hoc analyses are not available if they must compete with higher priority jobs for mainframe time. In some states and territories, quick turnaround analyses would still be done with sorts of hard-copy forms if it were not for Epi Info analyses.

The ease of adopting and using Epi Info has facilitated the transition to NETSS and made it worthwhile in the states and territories. The transition to Epi Info was a smooth **one for the** states and territories which use it and were **visited as a part of this** Project. **Almost anyone** can be taught to use Epi Info once it is programmed. The manual is **easy** to follow, although it was suggested more than once that it would be hard to learn to program from the manual. However, programming of Epi Info applications is within the reach of the moderately computer literate. Most state health department staff interviewed had written at least some Epi Info programs. **The** limitation on programming may have as much to do with statistical expertise as with the software **application**.

The improved access to data provided by Epi Info has made several states and territories easier to enter program-specific data for high priority diseases. Hepatitis was most commonly mentioned as a candidate for addition to NETSS in those states and territories visited for this project. State epidemiologists would like a better handle on hepatitis risk factors and on the number of carriers that would become available with better access to surveillance data. States and territories recognize the additional burden of data entry to add program-specific data to NETSS. However, **they feel that it is worth it** for improved access to data on high priority conditions. Many states already enter these data for their own use.

Epi Info provides a common software supporting data entry, data management, data transmission and data analysis across federal, state and local public health agencies. In particular, it provides a basis for distributed data entry in district and local health departments. Epi Info has been successfully taught to local staff by state personnel and is working effectively in three of the four states and territories visited with remote data entry. It is probable that the large number of distributed data entry initiatives encountered in this study would have been much slower to develop in the absence of a standard software package which is easy to learn and use.

### **Future Directions in Surveillance in States and Territories**

An important part of this study was an assessment of the directions in which surveillance is going at the state level. In this section, we suggest some trends for the future that have emerged from interviews with state public health staff.

State epidemiologists generally do not reject the idea of a standard implementation of NETSS, as long as it is flexible enough to accommodate their own needs. State staff interviewed for this project did not reject the idea of a more standard NETSS/Epi Info implementation out of hand, although several placed conditions on the acceptance of a standard package for ongoing surveillance. One state expressed a serious objection to use of a standard package because they do not have staff to replace the CDC support function and to expand the system to the local level. Most states and territories would find submission of a standard core record acceptable if they had the flexibility to expand the record to include data which they collect and use but do not send to CDC. States and territories do not want to lose what they have already developed for their own use. Interviewees in two states suggested that they would support this if it improved technical support from CDC.

Training and technical support of NETSS/Epi Info and the development of expertise in the system at the state level would both be easier, if there were a more standardized version of the software applications used in all states and territories. A more standard package would permit a single resolution of common problems for all states and territories and facilitate sharing of information among states, territories and localities. Revisions, upgrades and other technical support could be provided to states and territories on disk in a more timely fashion than is possible when CDC staff must travel to states. A

more standard package would also reduce the opportunity for errors because there would be fewer programs in use. The drawback to such a change is that states and territories would need to use a more standard reporting format than they presently have. Customized installation of NETSS and Epi Info for each state has been an important selling point in persuading states and territories to make the investment in NETSS.

There seems to be a trend toward a broadening of surveillance beyond infectious diseases.

Interest in developing and enhancing systems for monitoring chronic diseases, environmental illnesses and injuries were mentioned in several states. In two states, there are pilot projects ongoing to track non-communicable diseases. Even in states where there are no initiatives at present, this idea came up in conversation. This suggests that the demand in states and territories for the extended NETSS format to support surveillance data other than that for communicable disease may grow in the future.

Reliance of states and territories on CDC for training and technical support of Epi Info and NETSS is likely to persist until state epidemiology offices have developed their own expertise.

There will be a continued need for training and technical support of NETSS and Epi Info in states which have not yet completed the transition to this method of submitting surveillance data. At the same time, one or more individuals with a high level of computer competence were found in four of the states visited for this project. To an extent, state health departments must acquire their own expertise. The transition from mainframes to personal computers has not been accompanied by a corresponding shift in state staff available to support computer applications. Partially this is due to lack of staff because of budget limitations and hiring freezes. In many of the states and territories visited for this project, there were no staff positions in the state budget to accommodate support of personal computer applications. In one state visited, there is a state policy prohibiting state support of computer packages that are federally funded or developed. In others, the computer support staff that exist are not knowledgeable of NETSS and Epi Info. However, it was clear in this study that the level of computer sophistication among public health staff in the states is increasing rapidly.

Epi Info will become a standard for entry, management and analysis of infectious disease surveillance data in states and territories. All of the evidence from this study suggests that use of Epi Info by states and localities will expand in the future, especially with the full implementation of Epi Info, Version 5. Epi Info provides software support to both of the high priority concerns expressed by state epidemiologists. Even Minnesota and New York, the two non-Epi Info states chosen for this study, use it for some purposes. New York is using it for distributed data entry at the local level and is contemplating a switch to an Epi Info based system. Minnesota uses it for ad hoc analyses.

Distributed data entry will become a high priority in large and medium states and territories. Staff in the four medium and large states and territories visited for this project are pursuing the goal of data entry at the district or local level. This not only reduces data entry burden by decreasing the number of cases entered by any one agency, but will improve data quality by moving error resolution closer to the source of information. There is a size effect in the move toward distributed data entry. The two states with the smallest numbers of cases plan to retain control of data entry at the state level.

Connectivity with other states and territories and with other health agencies within the state is a low priority for state epidemiologists. Direct connectivity with other states and territories was explicitly identified as a non-priority by all but one state with a large number of neighbors. Most interviewees felt that telephone calls and facsimiles are adequate for exchange of data between states and territories. Computer links with other state agencies, such as public health laboratories and vital records offices, were also a low priority in all states and territories visited. However, Local Area Networks (LANS) are either operating or in the planning stages in four of six states and territories visited for this project. Computer staff in these states and territories are very enthusiastic about LANS and plan to extend them as widely as possible within the health department. Also, the priority of intra-state connectivity may increase if surveillance of environmental or chronic health problems increases in importance.

Improved quality control and better analysis capabilities are high priority concerns for future developments in surveillance at the state level. State epidemiologists were asked to rank their priorities for development among better quality control of data, enhanced analysis capability, greater connectivity with other states or with other agencies in their own states, and improved dissemination of surveillance data from CDC. Enhanced analysis capability and data quality control were each mentioned by two states and territories as the top priority. Improved dissemination of national data was not a clear priority, although there is widespread interest in being able to access and analyze data using WONDER. Computer links between state health departments and district or local health departments are the only connectivity which have immediate relevance for state epidemiology staff.

The analysis enhancements to Epi Info most desired by states and territories are new analysis capabilities, including custom mapping. State staff were asked what kinds of Epi Info enhancements they would like. Epi Map is eagerly anticipated in several states and likely to be well received, especially if states can generate customized maps. There is interest in pre-programmed Epi Info applications for common epidemiologic analyses such as frequency tables. Most people would like a pre-programmed Figure 1 capability for their own state. All states and territories visited either have or will soon accumulate the five years of baseline data needed to support such an analysis. Staff interviewed are very interested in the Epi Workstation concept. "Canned" analyses of common epidemiologic analyses, access to program management and budget data, and "tickler" systems to identify cases overdue for follow up investigation were suggested as possible modules in an Epi Workstation.

Enhancement of Epi Info graphics capability has a lower priority from the perspective of states and territories. It is likely that states will continue to use commercial graphics packages to produce camera-ready graphics. They are already accustomed to doing this, know the software, have the programs written and see little to be gained by substituting Epi Info graphics.

### **An Inventory of State Resources to Support Surveillance**

An important product of the state study was an inventory of hardware, software and staff resources available in states and territories to support the assessment function in the health department. This inventory is presented in tabular form in Table 5. This table provides an overall indication of the amount and kind of software found in the states and territories studied.

Most states and territories were found to be well-equipped with hardware and software. In several states and territories, local health departments had inherited computers from the state health department as the latter upgraded to larger and faster computers. Moreover, a high level of computer literacy exists in state health departments at professional and clerical levels in spite of a lack of specialized technical support. Most states and territories have courses in basic computer skills, word processing and DOS available either in the health department or in other agencies. State-provided training for local personnel is almost non-existent. Three of the states and territories visited have provided training in Epi Info for local personnel. These initiatives have been highly successful and should be emulated in other states and territories.

**TABLE 5.**  
**COMPUTER EQUIPMENT AVAILABLE**  
**TO NETSS AND DISEASE SURVEILLANCE**

	MO	NM	MN	SD	NY	PR
<b><u>Surveillance Coordinator</u></b>						
Type of Computer	Zenith 386 16 Mh	IBM AT 6 Mh	Epson AT 12 Mh	386 Clone 25 Mh	AST 386 33 Mh	IBM Model 55
Math Coprocessor?	NO	NO	YES	YES	NO	YES
Disk Drives	5", 3", HD	5", HD	5", 3", HD	5", 3", HD	5", 3", HD	3", HD
Hard Disk Drives	120 MB	20 MB	40 MB	40 MB	110 MB	30 MB
Tape Backup?	NO	NO	N	YES	NO	NO
Mouse?	NO	NO	YES	NO	NO	NO
Monitor	VGA	Mono w/Graphics	EGA	VGA	VGA	VGA
Modem	2400b	1200b	2400b	2400b	2400b	2400b
Printer	IBM Proprinter	Toshiba PI 351 w.c. d.m.p.	Laser *	Panasonic KX-P1524 w.c. d.m.p.	Epson PX-286E  w.c. d.m.p.	IBM Proprinter XLE
<b><u>Available at Local Offices</u></b>	Districts have XT's or 286 PC's	COMPAC Portable PC w/20 MB Hard Disk (Albuquerque)	**	**	Terminals in County Commissioner Offices.  PH Nurses have AST 286 PC's	Districts have IBM 286 PC's

\* Available through LAN

\*\* Single Data Entry Point

w.c. d.m.p = Wide carriage dot matrix printer

TABLE 5. (Continued)

**COMPUTER EQUIPMENT AVAILABLE  
TO NETSS AND DISEASE SURVEILLANCE**

	MO	NM	MN	SD	NY	PR
<u>Available through State Surveillance Office</u>						
Office Environment	Stand-Alone PC'S	Stand-Alone PC'S	SUN Etheret LAN	Dept-Wide LAN Token Ring, PC-LAN	3270 Emulation Boards for Mainframe Communication	Dept-Wide LAN
Additional Computers:						
XT					7	1
286	1	5	56		1	9
386	1	2	56	1	2	7
Laptop w/20 Mb HD				1		COMPAC Portable
MacIntosh	1		12			
Additional Printers:						
Dot Matrix Printers		1	2		Several	7
Color Plotters				1 *	1	2
Laser Printers	1	1	1	1	1	1
<u>Planned</u>	LAN for Section	2-PS/2 Computers Color Monitor				
<u>"Wish List":</u>	Multicolor Plotter					Optical Scanners

## CHAPTER 5. CONCLUSIONS AND RECOMMENDATIONS

In this chapter we bring together the findings from the CDC and state studies and draw some general conclusions about the present operation of NETSS and directions for electronic surveillance over the next few years. We summarize the hardware, software and staff resources available in states and territories. At the end of this chapter, we make recommendations for future support of surveillance based on the findings of this study.

### CONCLUSIONS

#### Standardization and Customization

Although customized implementation of NETSS was a selling point **for its adoption, state and territorial staff interviewed for this project did not reject the idea of a** standard implementation of NETSS. Persons interviewed generally understand the need for a standard format to guarantee comparability of data from different reporting jurisdictions and to permit better quality control. Some epidemiologists cited the year 2000 objective for a universal set of public health data. Others support the idea of greater standardization if it would lead to better technical support.

There is already a great deal of standardization in NETSS. All **states visited** use the **CDC/CSTE** case definitions. Most states are transmitting required “core” data to CDC over NETSS. In practice, there is not much variation from state to state in the kinds of data collected for notifiable diseases or in the subset of data which are transmitted to CDC. Variation in the precise diseases which are notifiable is somewhat greater, but this would be no more of a problem with a standard NETSS implementation than it is in the present one.

There were reservations expressed about standardization in states and territories. The standard implementation must be flexible enough to accommodate the needs of states. It must be possible to expand the state system to include variables and disease types which states must collect and use but which they do not send to CDC. States and territories do not want to lose what they have already developed for their own use. **EPO has covered this contingency** in their plan to build customized, state-specific screens as part of the standard implementation of NETSS.

Adequate support in the transition to the standard system from present procedures may be more of a concern to states than is maintenance of the system once

established. The one serious objection to a standard implementation of NETSS came from a concern that CDC might not be able to provide enough support of the transition to NETSS and Epi Info. Changes in procedures for managing surveillance data are risky for states and territories because they cannot bring down the system to accommodate transitions. The transition to Epi Info and NETSS has been easiest in states and territories which had no prior computerization of surveillance because nothing had to be disassembled to allow replacement by the new system. However, introduction of a standard implementation of NETSS would put all states in the position of making the transition from one computer system (their present customized version of NETSS) to another (the new standard version of **NETSS**). This will be least disruptive in states which already have forms and coding schemes approaching the standard version. It will be more of a problem in large states or in states which have not completed the transition to a PC-based system.

There is a separate issue of the feasibility of standardization across states and territories of records prepared in the extended NETSS format by state epidemiologists or other state and local public health staff. With a few exceptions in some states, **disease-specific** reports of infectious diseases to CDC use a nationally standardized form in all states. Development of standard NETSS modules for these diseases can be directly derived from existing forms. In fact, this kind of reporting may be easier to standardize than is general morbidity reporting which varies from state to state. New forms of reporting, such as that proposed for childhood lead toxicity, can be standardized from the outset. However, ongoing surveillance and/or registries of chronic conditions or environmental exposure to health risk may require customized development of NETSS modules if transmitted data are to be comparable to those pre-dating the automated data transmission. In addition, for data not presently under national surveillance, differences in the kind and amount of data collected may not permit use of a standard program for data entry and transmission.

#### Training and Technical Support

States and territories place heavy reliance on CDC for technical support of start-up, ongoing operation and staff training for NETSS and Epi Info. Up to the present time, EPO has been responsible for all activities related to the development and implementation of NETSS. This includes development of programs, installation of

programs in states and territories, training in the use of NETSS and Epi Info, and ongoing technical support of these developments. This has become a very large job.

Epi Info and NETSS are designed to be used with little difficulty by users with little or no previous computer experience. Generally speaking, the system operates this way in daily use. However, problems arise when something goes wrong. When this happens, state epidemiology offices do not always have access to staff with computer training and expertise to help them out. Reliance of states and territories on CDC for training and technical support of Epi Info and NETSS is likely to persist at least until state epidemiology offices have grown their own expertise in this specific software.

There is evidence that this is happening in some states. In this study, the health departments which are the most successful users of NETSS and Epi Info are those fortunate enough to have at least one staff person who is both knowledgeable about public health and has an understanding of how computers interface with users and data. These people may or may not have training in computers, but they understand them and they like them. They have been an asset in easing the shift to NETSS because they explore the limits of the software and come to understand it from the inside out. They are in a position to fix small problems before they become big ones. Lacking this kind of support, trial-and-error (often under pressure) becomes a common way of resolving software problems.

However, states making transitions from mainframes to PC-based data management, states experiencing recurrent difficulties with NETSS, and states with a demonstrable lack of intrastate technical support in PC-based systems need help in implementing NETSS beyond that which can be provided in a site visit. Problems are likely to appear in ongoing operation of NETSS' which do not surface in an initial test of the setup. Often these require trial-and-error troubleshooting which cannot be accomplished from a remote location at CDC. In difficult situations, there needs to be someone available on-site to handle problems at the time that they appear.

Training and technical support of **NETSS/Epi Info** and the development of expertise in the system at the state level will be easier with a more standardized version of the software. A more standard package will permit one-time resolution of common problems. Revisions, upgrades and other technical support could be provided to states and territories on disk in a more timely fashion than is possible when CDC staff must travel to states and territories. A more standard package would also reduce the

opportunity for errors because there would be fewer programs in use. However even a standard format for NETSS will not reduce the support burden to insignificance. Support of Epi Info and of state-specific modules will still be needed. There will be an increasing support burden at CDC if NETSS is expanded to other CDC programs.

### **Software Development**

CDC has tried to be sensitive to state needs in use and development of software. CDC has made a commitment to honor the option of state health departments and other agencies to use any hardware or software configuration that they choose as long as they can provide CDC with the standard categories and coding schemes needed to process the data. However, they argue that consistent user interfaces would decrease the burden on data entry personnel and analysts who need to control alternative schemes if multiple software is used. The results of this study suggest that multiple software may be a decreasing problem over time, at least in the offices of state and territorial epidemiologists who process notifiable disease surveillance data.

The proliferation of software in the offices of state epidemiologists was seldom mentioned as a problem in this evaluation. This may reflect the operation of a gravitational effect of Epi Info on computer management of surveillance data. This phenomenon occurs in software development when specific products come to be adopted as standards as much because they are widely used as for their technological characteristics. This is seen frequently in spreadsheets or word processing packages which dominant sectors of the business world because people need to exchange computer files.

This is not to imply that public health agencies at CDC and in states do not use other kinds of software to process and transmit surveillance data. They certainly do. However, Epi Info was used for routine management and analyses of notifiable disease information in all four of the health departments visited which have this software. Even Minnesota and New York, the two non-Epi Info states chosen for this study, use it for some purposes. It is especially significant that it is used in 'New York to support distributed data entry. An increase in this practice may push states further toward an Epi Info standard.

Epi Info has changed the way in which public health is done in the United States by providing states with access to analyzed data which was not possible when staff had to either sort hard-copy forms or wait in the queue for mainframe analyses. It

has provided public health professionals with the power to access and analyze their data in creative ways. Priorities expressed for further developments of Epi Info reflect this. (See Table 4.) Enhanced analysis capability is top priority. State and territorial epidemiologists are interested in new analysis capabilities, including custom mapping. Most people would like a programmed Figure 1 capability for their own state. Epi Info graphics capability has a lower priority from the perspective of states and territories. It is likely that states will continue to use commercial graphics packages to produce camera-ready graphics.

Adoption of Epi Info and/or NETSS may be slower in CDC programs which have made investments in other kinds of software. Use of dBase and RASH for measles data is a good example of this. The system developed around RASH has been tailored to meet both disease control and programmatic needs related to measles. It generates a number of pre-defined reports which are needed by the program. There is little motivation to change from a working system as long as state immunization program staff agree to use it. This does not mean that the Division of Immunization has no interest whatsoever in NETSS. They are open to the possibility of using it for other immunizable disease such as rubella and pertussis.

Whatever kind of software interfaces CDC develops for introduction to Centers, Institutes and Offices at CDC or for use in states and territories, all components of the system must be adequately tested and demonstrated to function correctly before they are distributed. This includes hardware, software and modes of telecommunication. Not only must software be beta-tested, but applications of the software written for specific users must be tested and debugged on simulated data. Inadequate testing and premature implementation of systems has a high cost in lost confidence of potential users. It is also costly to those who must cope with failure of essential public health functions because of malfunctions in computer systems.

#### Telecommunications

Future NETSS developments depend on concomitant developments in IRMO. Direct transmission of data to CDC, development of timely transmission summaries and rapid dissemination of the data over WONDER depend on an effective communications gateway with the CDC communications computer. Both the feasibility of these innovations and their timing is related to the development of a two-way capability in WONDER by IRMO. Rapid turnaround of transmission summaries and analyses will not be

possible until the ADATABASE/NATURAL rewrite of the NETSS data set is completed. It is essential that EPO and IRMO maintain communication and cooperation as these developments proceed.

### **Data Exchange/Connectivity**

Distributed data entry i.e. entry of data in district and local health departments rather than in the state health department is a high priority in medium and large states where data entry is a heavy burden. Plans for the transfer of program-specific data to CDC using an extended NETSS record also depend on remote entry in clinics, hospitals and other agencies outside of the office of the state epidemiologists. From the perspective of states, distributed data entry not only reduces data entry burden by decreasing the number of cases entered by any one agency, it improves data quality by moving error resolution closer to the source of information. There is a size effect in the move toward distributed data entry. Small states have no interest in it and plan to retain control of data entry at the state level. This suggests that distributed data entry is a response to the burden of data entry for a large number of cases rather than a desire for local quality control. However, staff in two of the four health departments with remote data entry emphasized that they seek to create a sense of local ownership of the data.

A problem with distributed data entry as it is now implemented is the **two-tiered** structure it introduces to error resolution. EPO must recognize that an error identified during a transmission from the state may require resolution at the county level. Programming at CDC that requires immediate error resolution will cause bottlenecks for the busiest states. On the other hand, delays in identifying errors to the state will be compounded by the time taken to communicate them to the appropriate county, presenting the possibility of three separate **datasets** existing at any given time and adding to the confusion. Allowing simultaneous transmissions by the counties to both CDC and the state epidemiology office would introduce other complications. Not the least of these is requiring staff from large states to make themselves available for the transmissions of as many as fifty counties. More useful would be the placement of error reports under the contributor's account that can be addressed periodically by the states and referred by them to the county.

Distributed data entry does not necessarily imply a distributed data base model i.e. a network of data bases linked by software in such a way that they appear to the user

as a single data base. A distributed database model may be applied to processing of data at CDC once it has been transmitted through a single entry point, but is applicable to the nation-wide NETSS in only a general way. States and territories taken together resemble a distributed database in the independent management of each state's respective data. Because nation-wide surveillance must acknowledge the responsibilities and prerogatives of the states, however, making data from different sites universally available throughout is not only impractical but inappropriate. Also, building a distributed database is a commitment of resources to accommodating different hardware and operating systems, while EPO's position is better served with standardization. Since the relationship between states and counties or districts is different from that between states and CDC, a distributed database model could be applied to distributed data entry at the county or district level. The limited software presently available that supports distributed databases are felt to be better at reading from databases than writing to them. These are long term options that states should consider as they develop distributed data entry at county or district level.

Many of the functions of distributed databases are being assumed by IRMO's implementation of WONDER on the CDC mainframe. WONDER is expected to distribute program data that is received through a communications gateway, allowing program managers control of their data while the mechanics of database management are handled by the system. This is preferable to having EPO assume the burden of distributing program data that is transmitted over NETSS.

Connectivity with other states and territories and with other health agencies within the state is a low priority for state epidemiologists. Most interviewees felt that telephone calls and facsimiles are adequate **for exchange of data between states and territories. Computer links with other state agencies, such as public health laboratories and vital records offices, was also a low priority** in all states and territories visited. In only one state was there a strong objection to independent transfer of data from agencies within the state to CDC. Most states felt that simultaneous transmission would be acceptable under most conditions. **However, transmission of communicable disease data** must be done with the knowledge and consent of the state or territorial epidemiologist.

The issue of the gold standard raised by the SCG committee was not specifically addressed in this project. **However, it is an important enough issue** to bear some consideration in light of what has been learned here. Where multiple data sets exist

for the same provisional data (such as those maintained at the federal, state and local level for surveillance data), the “gold standard” is that data set which is assumed to be the most accurate against which others are verified. The gold standard must be designated at the local, state/territorial or federal level. The set chosen should be one which is accessible to all users and is the most accurate available.

These characteristics do not necessarily belong to a single data set. The surveillance data set most accessible to all users would be the federal data set accessed through WONDER. The one which will receive the first notice of changes in the status of cases is the local data. The legal and certainly the most politically acceptable standard would be the state data base. State epidemiologists have legal responsibility for disease control and have the greatest vested interest in assuring the accuracy of the data. They will find and resolve inconsistencies in a timely and scrupulous manner because they must. The idea that the gold standard should reside on the CDC computer is likely to be a sensitive one for state epidemiologists. If CDC wishes to argue this, they should be meticulous in getting buy-in from CSTE.

### **Dissemination**

The usefulness of surveillance, especially infectious disease surveillance, is improved by the rapid collection and dissemination of data made possible by computer transmission. Not only is the initial assembly of data faster, but updating can be performed more rapidly by transmission of corrections. This improves the quality of guidance provided to state health officials by their own data.

Access to the two-way WONDER capability and completion of the move of the NETSS data set to disk will simplify the job of state epidemiologists by providing timely transmission and/or reconciliation summaries for NETSS transmissions. Production of these reports immediately after the NETSS transmission would facilitate ongoing reconciliation of state and CDC data sets and avert the annual reconciliation which can take up to eight months after the end of the year to complete. They would also improve data quality by identifying questionable cases while they are still cases and follow-up is possible.

However, rapid turnaround of data is not necessarily the first priority for all surveillance data. National surveillance is neither timely enough to support the control of infectious disease at the state level nor is it needed for this purpose. In this evaluation,

and in the previous one (2), we have found that even weekly provisional MMWR data are not timely enough to support action by state epidemiologists. Perhaps more importantly, since states provide this data and update it, their information is as current or more current than that published in the MMWR. This would not change if the MMWR data were put up on WONDER and could be accessed immediately.

The rapid turnaround of data may be less important to other CDC programs than is better quality control and analysis capability. Accuracy and completeness of data quality may be more important than timeliness for chronic and environmental diseases and even for laboratory data.

Improved dissemination of data through WONDER raises issues of confidentiality and data ownership. The technology exists to protect access to data but systems such as chronic disease registries may need to maintain data for follow-up that should not be disseminated. Small jurisdictions, whether they are states or localities, may have strong concerns with small cell size. South Dakota is an example of a state with legal protection of confidentiality which may impact on the kinds of data that they can report.

#### **Hardware and Software Resources in States and Territories**

Epidemiology offices in the states and territories visited for this project have or will soon have the hardware necessary to support present and planned enhancements to NETSS. All of the state and territorial health departments visited have computers with at least a 286 processor. Three of them have 386 processors. Disk space is adequate to run the Epi Info software, although storage of year-to-date NETSS data may become a problem in states with a very large number of cases. As hardware prices fall, the actual cost of upgrading equipment may be less than the perceived disruption and risk involved in moving the system. Better understanding of the programming they use should make hardware upgrades less difficult. The data transmission hardware and software needed for a direct link to the CDC mainframe is already present in the states and territories and is used successfully to send NETSS transmissions to DIAL COM. Most counties and districts also had transmission hardware and software with which to transmit locally entered data. The most pressing need in states and territories was for NETSS-related training and technical support.

## RECOMMENDATIONS

### Standardization and Customization

The issue of standardization versus customization must be considered from a local perspective. Standardization is a matter of degree, and without details makes little sense from the point of view of a single participant. Despite the customized installations, the adoption of Epi Info by 36 states and territories in itself constitutes a degree of standardization, and this was often the view of state representatives. States that have abandoned earlier systems in favor of Epi Info will not understand the issue of standardization versus customization unless the choices involved are presented to them clearly. Their perceptions of these choices will be crucial in obtaining support and coordinating their efforts.

The central role of EPO in choosing directions for standardization and coordinating that movement must also be recognized. EPO is regarded as the 'front office' for surveillance data by a network of states and territories contributing data on essentially a voluntary basis. The impetus behind that network is shared public health goals, confidence and credibility. EPO needs their cooperation and support to perform its mission, and is in the best position to avoid disruptions and misunderstandings.

CDC should proceed with development of a more standard NETSS package. Developing and implementing a standard NETSS package will involve three steps:

- Describing the NETSS system as it is presently implemented in all states and territories to the participants themselves,
- Defining the standard to be adopted and communicating that standard to the states and territories, and
- Implementing that standard incrementally with short-term benchmarks.

Describing the NETSS system as it presently exists will be necessary to show the states a common ground they can begin moving towards. The documentation that has been prepared and has been available to this review does not adequately describe the efforts EPO has expended in customizing installations to date, neither to the states that are dependent upon those installations nor to EPO itself. EPO should begin work on compiling

the following as part of a full documentation that can be made available to the states and territories:

- A comprehensive list of the diseases under surveillance and the variables collected for them should be compiled for each state and territory. From these lists comparative tables of diseases and variables by state can be drawn.
- Reporting forms in use in states and territories should be collected.
- Prints of the data input screens prepared for states and territories should be collected and reviewed.
- A comprehensive list of Epi Info Check programs that are in use should be compiled. This will allow comparative information on permissible values and edit checks for variables in common usage among states. Whether inconsistencies are uncovered or consistent editing is confirmed, the exercise will provide valuable information and illustrate the value of further standardization.
- A comprehensive list of Epi Info Analysis programs that are in use should also be compiled, including but not limited to those written by EPO staff. Reports created by these programs should also be collected.

The first three items, if not already available, should not be difficult to request of the states. Compiling comparative tables should be possible with 80 to 120 hours of clerical support. Appropriate Check and Analysis programs should be easily located with the help of EPO staff over the phone. Assembling a table that presents edits and ranges for comparable variables across states could be a daunting task considering the detail supported by Check programs. However, it is unclear how a standard package could be developed that did not take into consideration the way variables are presently edited, so the work would not be wasted. Also, since the Check programs themselves are ASCII files that follow a standard block format, software could be developed that reads and documents the Check programs themselves. Such software could be extended to compare practices among states and generate documentation for the entire system.

With the completion of comprehensive documentation, work should begin on the development of prototype standard reporting forms that states can review. The assembled documentation can serve to show the states how they differ in their uses of 'Epi Info and NETSS so the breadth of current customization and the possible advantages to

standardization become apparent. Commonalities in the variables reported and the way they are edited, standard reports and analyses, and similarities in reporting forms should emerge that will suggest directions for a standard. Distributing the documentation while soliciting input on a standard will enhance credibility and confidence throughout.

A standard reporting form and NETSS installation will need promotion and advertising. Enticements such as an Epi info program to generate a Figure I-type report for the standard, accompanied by a newsletter article explaining how it works and how it could be adapted to non-standard NETSS installations could encourage interest in the standard.

Implementation of the standard must be coordinated with incremental benchmarks and a schedule for the entire program of development. At the present time, multiple planned enhancements of NETSS are being developed without enough attention to their linkages to one another. This leads to problems with timing which impede the efficiency of NETSS development and results in delays and additional development costs. For example, the **ADABASE** re-write is made more complex than it need be by the need to accommodate both the 30 byte and the 60 byte core record format. Direct transmission to CDC, the availability of rapid transmission summaries and analyses of data all depend on the **ADABASE** natural rewrite. We make specific recommendations about developmental priorities in the last section of this chapter. Here we are suggesting only that a plan must be made and accepted by all persons involved in the program to enhance NETSS.

#### Training and Technical Support

EPO has provided extensive training and technical support in the adoption and use of NETSS from its inception. Continued support is implicit in getting the states to participate in any of the planned changes to NETSS or in establishing a standard NETSS installation. States cannot be expected to make changes and risk disruptions in a vacuum. The issue is what form that support should take.

Providing a disk with an installation program and no further assistance or explanation is not adequate support for the kinds of major revisions envisioned by EPO. This may be an acceptable approach for commercial software that is actively sought and purchased, but is a poor way to encourage confidence in modifications to ongoing operations that states rely

upon. Our evaluation at the states found two occasions where disks like these were simply passed on to another office or ignored.

EPO should be prepared to provide assistance in installing and using software sent directly from IRMO to states if no other support is available. As said above, the provision of floppy disks to install upgrades is inadequate and potentially disruptive. However if that is the limit of the support that is available for CDC-based surveillance software, then it is in EPO's own interest to provide assistance where it can, even if it means helping states to install such products and providing clarification on their use. EPO's central role in surveillance is a two way street that should be recognized by all concerned. By serving as liaison between the states and offices such as IRMO, problems in software can at least be identified and documented. IRMO gains information on such products as WONDER from the field, EPO gains control over potential disruptions and additional resources, and the states gain a centrally positioned advocate. This is an expansion of EPO's role that would require a commensurate staff increase of at least one FTE.

This approach is preferable to IRMO's practice of providing support through an outside contractor. New faces in themselves can be a disruptive influence in busy state health departments. Outside contractors are less likely to encourage the ongoing relationship between support staff and state staff which EPO has already established. Finally, CDC must pay for a learning curve of contractor staff who will need to acquire familiarity with disease surveillance systems.

EPO should continue its on-site installation of Epi info and NETSS in states and territories for the immediate future but should focus on capacity-building in states as a long-range goal. We concur with the recommendation of the SCG on this issue. Public health assessment is an essential public health function that should not be put at risk by delays and/or errors in the installation of software. A support model which relies on manuals and tutorials with little on-site support will not be appropriate for NETSS in the near future. However, long-term support of ongoing surveillance is beyond CDC's mission in state-based surveillance and is very consuming of task time and funds. The long-range goal of enabling states to develop and strengthen their own computer capabilities should be integral to EPO planning for future NETSS development. This will be done by moving to a standard NETSS system, by providing states with better documentation of their own

NETSS system, and fostering the development of computer expertise in state health departments.

Training and technical support should rely less on "turnkey" systems and seek to educate the states on the system they are using. Most of the Epi Info installations we looked at included DOS batch files, PROCOMM script files, and Epi Info Check'and Analysis program files in addition to the data entry screens, data files and the Epi Info software itself. State personnel were not always clear which programs did what and approached their systems gingerly and apprehensively. Problems that arise become magnified because the entire system is viewed as somewhat mysterious, and identifying and communicating problems becomes much more difficult. Nonetheless state staff report using a wide range of software packages along with Epi Info such as SAS, Harvard Graphics, and others. Also, representatives of a non-Epi Info state expressed reluctance to become dependent upon EPO for programming support. The use of batch files to create menus and run Epi Info modules from DOS is useful, but should be accompanied by documentation and fuller explanation of the installation.

Documentation of each installation should include the following:

- A list of all DOS batch files, the files or commands that call them and the files they in turn execute, along with their overall function.
- A list of communications script files that are used.
- A list of Epi Info .EPA, .PGM, .CHK and .RPT files, their function and the files they read.
- A table documenting the ranges and edits programmed with .CHK files.
- A list of the .QES and .REC files that are generated.
- A flow chart describing the system, the sequence involved in generating the NETSS transmission and their own database, and the subdirectories on which they reside.
- A list of likely error messages and their source (DOS, Epi Info) that can help identify and communicate problems.

Providing such documentation will increase the workload in site visits, and EPO staff are spread thin already. Over the long run, however, comprehensive documentation of installations will reduce the need for further site visits and will allow more problems to be resolved over the phone or with carbon copy sessions. As states move toward a standard NETSS installation, the documentation requirements will diminish. The documentation will encourage movement towards a standard, as states see exactly where they are and what they are buying into.

This is not to suggest that all staff at state levels can be educated in the technical aspects of Epi Info and NETSS. Public health nurses are often quite unfamiliar with computers and have little time to develop the skill. However a failure to document programs and files as they have been installed at state levels insures that no one there will develop an independent familiarity with their own system.

CDC should create a computer development internship to provide long-term computer expertise to states with special needs. An intern program would permit computer-literate scientists in the early part of their careers to support states with very large caseloads, states making transitions from mainframes to PC-based data management, states experiencing recurrent difficulties with NETSS, and states with a demonstrable lack of intrastate technical support in PC-based systems. Interns should be assigned to states for a period of 90 days to one year to provide on-site technical support in operational phases of NETSS operation and other aspects of computerized surveillance. In states, they should be responsible for installation of the standard NETSS package, training of state staff in its operation, routine troubleshooting, and briefing CDC staff on unmet state needs. They should also prepare comprehensive documentation of the state installation according to a standard documentation format.

The model for interns is taken from the computer literate staff member who provides daily support to Epi Info and NETSS in some of the states visited **for this study**. **CDC can support this capability in states which have not been fortunate enough to find** such a person on their staff. We are not recommending a permanent commitment of CDC staff to support state computer operations in individual states. An intern approach is recommended because both the need for this type of support and the length of an internship are limited. This would provide an excellent opportunity for public health staff with computer skills to initiate their careers with a state health department experience.

The most important criterion for selecting interns should be demonstrated capability and expertise in computer applications and telecommunications. The best candidates for interns are public health staff at the beginning of their careers who have a bachelor's or Master's degree in a health or statistics related field. Computer scientists are not to be ruled out, but may not make the best candidates for this position. Computer scientists tend to have specialized expertise in advanced computer applications. These individuals may not be able or willing to maintain the kind of communication with non-computer specialists required by this job.

CDC staff experienced in the installation of NETSS in states should participate in the selection of interns, train and supervise them, coordinate their activities in the states and monitor progress of the states and of the intern program. Interns should be trained in Epi Info and NETSS for a period of at least six weeks. Training should include an orientation in IRMO operations that impact on surveillance and hands-on experience in a state which has a well-functioning NETSS and Epi Info system.

CDC training in Epi Info should focus on training trainers to provide support to local health departments in the states. This follows from the previous recommendation. Less emphasis on a turnkey approach and better documentation of installations will build tools that EPO can use to maximize the return on training investments by developing courses, manuals and materials to support in-state training by state or local health department staff. The training of staff in thousands of local health departments in state-specific surveillance procedures is beyond the ability of EPO staff to conduct, and the installations themselves could become impossible to maintain. EPO should work to build local support that is possible by clear documentation and evolving standards.

CDC should promote the teaching and use of Epi Info software in Schools of Public Health and other forums. Many of the staff in Puerto Rico know Epi Info, Version 3, before they came to work at the health department because it was part of their Masters of Public Health Curriculum at the University of Puerto Rico. CDC should actively promote the widespread teaching of Epi Info by sending brochures and/or disks to those responsible for curriculum development for M.P.H. curricula and other public health courses. This should be done in much the same way that textbooks and other teaching materials are promoted. Although it will take several years for the benefits of this strategy to materialize in state

health departments, an investment of several thousand dollars could eventually reduce the support burden to states to occasional support of unusual problems. Epi Info should also be placed in university libraries, bulletin boards and in public domain software houses.

CDC should support development of a mechanism for exchange of NETSS and Epi Info support among states. State staff who have developed expertise in the operation of NETSS and Epi Info have a perspective on the system which may not be easy to find at CDC. Staff members at states and territories share common experiences that can be a resource to everyone. A problem in one state may already have been solved in another. Distribution of system-wide documentation can serve to point out similarities in processing and help the states to identify others with whom to exchange ideas and information. A newsletter can also circulate information, post questions and notices by immediately using resources already in place. A newsletter can immediately be included with the Transmission Summary Reports now mailed on a weekly basis or posted as sign-on messages and "READ.ME" files at DIALCOM and later on the CDC network.

#### Software Development

New or enhanced software developments should not be introduced into states and territories until both the software itself and specific applications have been well tested.

Admittedly, it is not possible to completely prevent malfunctions in the field. Further, if EPO assumes a liaison role for software produced by IRMO, this will to some extent be out of their hands. However, CDC should be very conservative in the promises they make or imply about the performance of software. Put as simply as possible, all software must work where and when it is applied in the field. Many state staff who use Epi info and NETSS do not understand the reasons for malfunctioning software and cannot distinguish errors in the Epi Info program itself from conditions that were not anticipated in the Check and Analysis programs that were written for them. Better documentation at the installations, greater familiarity with the system and further standardization will help in the long run, but the immediate problem for the state is a costly disruption. Historical data can provide a means of comparing software upgrades with the records as they were processed by the previous system. Simulated data can be useful as well, however the simulated data should be sufficiently realistic. Any testing of new software should look closely at the generation and processing of unique ID numbers.

Resources should not be devoted to competing with commercially available software that may already be in use in the states. State personnel are using a wide range of commercial software along with Epi info, such as SAS, Harvard Graphics and desktop publishing packages. A more fruitful direction may be in expanding the capabilities of Epi Info's Import and Convert modules to accommodate specific programs. While states often felt that the charts and graphs available through Epi Info were not acceptable for publication or presentation, the limitation posed no practical problem because of other software resources. This may not be the case with all states or localities, however, and commercial software may present a substantial investment to some. But to many others upgrading Epi Info to provide statistical procedures comparable to one package and graphic quality comparable to another presents a limited net gain in resources. An exception to this is Epi Map, which enjoys a strong interest in the states. Mapping software is less commonly available in the states or imposes excessive hard disk space or processing requirements.

Modifications to Epi Info to accommodate new operating systems and environments should be considered on a case by case basis. Devoting resources to adapt Epi Info to new operating systems and environments should be considered carefully. It would serve little purpose to develop an Epi Info version for Windows at a time when several large states are busy coordinating local data entry and would probably prefer that installations remain simple. Also, hardware at the local level is generally available and adequate for the current version of Epi Info but may not be able to use further advances. Windows requires a 386 processor to use it for anything more than a graphics interface, and generally needs a great deal of hard disk space, 25mh processor speed and 4 megabytes or more RAM to take advantage of its multi-tasking capabilities. This may become an interest to large states in the future, particularly in handling transmissions from dozens of counties, but will be a low priority for the immediate future. Development of a graphical user interface would similarly have a limited audience, since only one of the six PC's used by surveillance operators had a mouse attached. No special accommodation for DOS 5.0 should be necessary, other than that additional memory may be available and that operators may want to run their systems from the DOS Shell provided with it.

A fully LAN-Compatible EPI Info should be given a high priority. Four of the six states and territories visited have or are installing LANs in the health department. Although two of the states with LANs intend to maintain a centralized surveillance system, LANs may present further opportunities for distributed data entry. Compatibility for DOS-based networks should aim for high level LAN systems such as NOVELL 386 or NOVELL SFT that can support file access protection and a communications gateway. It should be noted that one state visited expressed concerns about placing protecting the confidentiality of surveillance data placed on a network. States were very interested in EPI Info with a multi-user capability.

Epi Info enhancements should serve data management needs or specific public health objectives. Enhancements should look to Epi Info as it will be used to support a distributed data entry function in large states. At present it cannot easily accommodate a large number of valid hospital or town names and distinguish their applicability to individual counties. Separate installations for each county could be used but would require considerable support for just a single state. The ability to access a table look-up system could streamline installation in many states in the future. The Validate module should be enhanced to allow the validation of only priority variables rather than the entire record. Value and variable labelling could also be improved for Analysis tabulations. There is also interest in the ability to log and track work sessions and keystrokes for budget reports.

#### Telecommunications

The development of a two-way communication capability between states and the CDC mainframe, such as that proposed for the new implementation of WONDER, is needed to support expansion of the NETSS system. The extension of NETSS to other programs, the management of accumulating NNDSS data, production of rapid transmission summaries, and direct transmission of NETSS data to CDC all require such an implementation. Development of NETSS on EPO and of WONDER in IRMO must be coordinated. We recognize that NETSS and WONDER are both far along in their development, and that some of the recommendations which follow would have been more effective if implemented earlier in the developmental process of both systems. However, we believe that the transition to linkage of 'NETSS to the CDC mainframe through PC WONDER can still be facilitated by the following measures.

CDC should designate a team including representation from EPO and IRMO to oversee development of a telecommunications aatewav between the CDC mainframe and remote users of NETSS. This team must assure that communications are clear enough so that the priorities of EPO and IRMO are consistent. If one of these groups must slow down or change developmental priority to take into account progress of the other, there must be a mechanism for making this kind of decision. We realize that such a team has already been designated, but wish to emphasize how critical we feel the effective operation of this team is to the future of NETSS.

The EPO/IRMO team should review and update the schedule and milestones for the telecommunications aatewav at least twice a year. This will serve to prevent slippage of schedules which is unknown to some members of the team. When such slippages occur, unrealistic projections may encourage premature commitments to manage data for users in states and at CDC. The schedule and milestones should be realistic and should err on the side of more time rather than less if there is doubt about the timely completion of some step in the process.

There should be a joint beta test of NETSS and WONDER as soon as possible and this should be repeated at strategic points in the development of both systems. If the movement of surveillance data in and out of the CDC mainframe through the WONDER "gateway" is to be done effectively, NETSS and WONDER must work well together. It is not adequate to demonstrate independently that NETSS works and that WONDER works. The earlier and the more frequently joint tests are conducted, the less the likelihood that a completed NETSS innovation will be found not to work correctly with WONDER or vice versa. Implementation of PC WONDER as a mechanism for capturing NETSS data should not occur until there has been a beta test of PC WONDER used for this specific application. States should not be used as beta test sites without their knowledge and consent.

EPO should actively assist IRMO in developing the "aatewav" to the CDC mainframe computer by ovoidina technical assistance to states in the use of PC WONDER. EPO has an important mission to support epidemiology and surveillance in state health departments which is not a part of the experience of IRMO. They have experience and knowledge of

the operation of surveillance in state public health programs which can help IRMO to successfully implement PC WONDER in the states. EPO is also likely to feel any negative repercussions of IRMO initiatives which operationally fail to perform in states. It is in the interest of EPO to do everything in its power to assure the success of PC WONDER.

Concrete actions which EPO can take might include pilot testing IRMO software in-house to see how it interacts with Epi Info or NETSS, alerting state epidemiologists to IRMO developments which are released or about to be released, and providing state staff with information on how to use IRMO software in site visits, regular training seminars, written materials sent to the states and telephone consultations. All of this will require ongoing cooperation and communication between EPO and IRMO staff working on the development of PC WONDER.

#### **Data Exchange/Connectivity**

Caution should be exercised in enlisting other CDC programs to adopt NETSS for their transmission. Agreements between program offices and EPO should be clearly defined and enumerated. Program data such as hepatitis and meningitis presents opportunity for the states to obtain data that they need and serves a valid surveillance function. Other program data may be suitable for NETSS transmission only with clear limits on the support EPO can provide. The suitability of datasets that are candidates for NETSS transmission should be evaluated in terms of the following:

- The data should include no identifying information for any purpose. If identifying information is required by the program office for processing at CDC, the data should be disallowed.
- Records should represent, on some level, a unique observation that can be uniquely identified with a data field comparable to the NETSS ID field.
- The number of records transmitted should increase total weekly NETSS volume by no more than 25%.
- Records should be limited to a fixed record length format without a compressed or packed decimal fields.
- Record formats should not be subject to change from year to year.
- Program data should not introduce an entirely new cadre of state staff to disease surveillance that will require further training.

- Adopting program data should present some advantage to the states. NETSS should not be adopted to avoid funding outlays.
- Indirect costs incurred by the states should be offset by the provision of edited **datasets** or reports.

CDC initiatives to encourage distributed data entry of infectious disease data in local health departments should be continued. This is clearly a high priority in medium and large states, where the reporting burden of some counties rivals many states. The hardware to support this will be available in most localities over the next few years, however this expanding circle of participants will be better served by simple, reliable software than by complicated installations requiring top-of-the-line equipment.

Inter-state and direct intra-state connectivity is a low priority for states and territories visited in this evaluation, and should not be a high priority for CDC. Existing methods for exchanging data within and between states are adequate for state needs. In addition, development of the telecommunications link with the CDC computer will provide states with access to data from other states without the need for a direct state-to-state link.

#### Dissemination

The most useful rapid turnaround summaries for states and territories would be line-listed reports of transmissions which would permit ongoing reconciliation of data sets. These could be used to resolve problems immediately while the case still exists and follow-up is possible. Present aggregate transmission summaries cannot be used for this because it is not possible to identify the cases which are in disagreement.

#### Developmental Priorities

There are a host of upgrades and enhancements to NETSS in various stages of planning and implementation at EPO which need to be prioritized. Many of the enhancements currently under consideration are contingent upon other upgrades, but a workable sequence for implementation has proven elusive. This has led EPO to a strategy of comprehensively upgrading single states in site visits, bringing the entire system into common usage over a projected two year period. Concurrent with this commitment is a complete re-write of computer processing for NETSS and ongoing development of a

communications gateway. Throughout all of this are long-term discussions of issues such as standardization that in reality have very immediate implications. It is necessary to distinguish changes and enhancements that can proceed independently, changes that can cause future bottlenecks, and changes that have immediate impact, and prioritize accordingly. We recommend the following priorities.

**First Priority:** The ADABASE/NATURAL re-write of the NETSS system at CDC should be **completed**. Whether states begin using the 60 character core record or the relational capabilities of Epi Info version 5 in two days or two years, their transmissions to CDC will continue uninterrupted. As a minimum, CDC must remain prepared to receive transmissions, process and disseminate data, and return error listings to the states as they have been doing. Failure in that will undermine all else. The PC Wonder Gateway is a central feature of present planning but is the third proposed replacement for DIALCOM, which has at least proven reliable and can be used during the transition to a better alternative. It will be much harder to retreat to the ad hoc programming that was used to maintain the NNDSS database if the **ADABASE** re-write stalls.

The **ADABASE** re-write addresses present and past problems such as overnight turnaround times and tape problems, and fits in with all the enhancements that have been discussed by supporting direct access to the data. Pursuing other changes without a processing system in place will be like driving a car with no front seat. The re-write has been complicated with the need to accommodate other changes, like accepting both 30 byte and 60 byte core records while states convert. This is part of the cost of designing a system to support future growth while being immediately usable. However, it has also suffered from changes in directions for telecommunications and the environment in which it will operate. The way in which data is to be passed to it and reports are returned must be defined and committed to. EPO should work to see that specifications for the systems interface are clear and that no further changes by **IRMO** or another party is allowed. **An FTE ADABASE programmer should be made available to EPO's current programming staff to support a sustained, full-time effort without interruptions or changes in direction.**

**Second Priority:** Detailed documentation of the NETSS system as it is presently installed in the states and territories should be EPO's second priority. A review of the NETSS installation in all the states and territories will point out directions for standardization and suggest areas where support and training requirements can be reduced.

**Third Priority:** The degree to which state installations can be standardized, from the use of common Epi Info programming to a standard reporting form, should be determined and a standard developed. Standardization should be either pursued or dismissed as an issue before further enhancements at the state level. With 36 states already using Epi Info, proceeding with further enhancements to customized installations will leave the issue largely moot and EPO's resources exhausted. EPO should first use its central position to develop and promote standardization among the states. This should be done in close coordination with the Surveillance Committee of the CSTE, and with ongoing input from the states and territories.

**Fourth Priority:** The standard implementation of NETSS should be introduced to states and territories. A standard installation should be defined to include Epi Info 5.01 and the 60 character core record format. The 60 character format is a high priority of the states and territories. It is felt to resolve problems with the five-byte ID field and the confirmation field, count field for reporting aggregate records, and the use of verification fields are all welcomed. (Of course, the use of these capabilities will depend upon the **ADABASE** re-write cited above.) A further incentive can be the inclusion of Epi Map if it has been adequately tested with real data.

**Fifth Priority:** EPO should do everything possible to promote an effective telecommunications "gateway" between states and CDC. The communications link that will ultimately handle all NETSS transmissions is essentially an **IRMO** responsibility but must remain an EPO priority as well. EPO is in a position to help make the telecommunications at CDC work by providing assistance in its use at the states. It should also work to remain in a position to insist that any telecommunications packages be tested thoroughly and comprehensively before states and territories are obliged to convert to it.

Sixth Priority: CDC should implement measures to provide adequate technical support to States seeking distributed data entry at the county level. Large states looking to adapt Epi Info for a distributed data entry system should be given high priority. However EPO support should wait to include the resources and opportunities that standardization can offer. Support for distributed data entry can be provided by state staff who have been trained by CDC to train local staff. In especially difficult or complex situations, states can apply for a computer development intern to support distributed data entry.

Ongoing: Software development priorities should be established on the basis of compatibility with the standard NETSS implementation and should not compete for resources with activities needed to maintain surveillance. Developments like multi-user capability and LAN compatibility should be given high priority in anticipation of the continued interest of states in distributed data entry. Also, developments to Epi Info that can enhance its usefulness in data management functions and improve data quality and efficiency throughout the whole system should continue. These developments are not dependent on other developments that EPO is pursuing. They can be conducted independently of other EPO priorities so long as they do not compete for resources of other EPO activities needed to maintain the public health assessment function conducted over NETSS. All subsequent Epi Info versions should be compatible with the standard installation of NETSS. Enhancements which will require significant modification of the standard NETSS implementation should be avoided.

## REFERENCES

1. Dull **H.B.** and Simon DG. Epidemiologic Surveillance Project Assessment and Evaluation, Final Report. Preventive Medicine Associates, September, 1986.
2. Odell M., Raichart D., Hersey J. and Arnold L. Final Report: An Evaluation of Infectious Disease Surveillance at CDC and in Five States. Battelle, May, 1990.
3. Report of the Surveillance Coordinating Group Subcommittee on Electronic Systems for Public Health Surveillance. CDC/ATSDR Surveillance Coordinating Group, January 18, 1991.
4. Wharton M., Chorba TL, Vogt R, Morse D, Buehler JW. Case Definitions for Public Health Surveillance. Morbidity and Mortality Weekly Report, Vol.39, RR-13. October 19, 1990.

**A P P E N D I X   A .**  
**P E R S O N S   C O N T A C T E D   F O R   T H I S   S T U D Y**

## APPENDIX A.

### PERSONS CONTACTED FOR THIS STUDY

#### I. AT CDC

Pat Coleman	Hepatitis Branch, Division of Viral and Rickettsial Diseases, CID
Andy Dean	Division of Surveillance and Epidemiology, EPO
Don Eddins	Division of Immunizations, CPS
Robert Fagan	Division of Surveillance and Epidemiology, EPO
Ron Fichtner	Division of Nutrition, CCDPHP
Emma Frazier	Behavioral Risk Factor Surveillance Branch, CCDPHP
Andy Friede	Information Resources and Management Office
Norma Gibbs	Division of Surveillance and Epidemiology, EPO
Joel Greenspan	Division of Sexually Transmitted Disease, CPS
Douglas Klaucke	Division of Surveillance and Epidemiology, EPO
Carol Knowles	Division of Surveillance and Epidemiology, EPO
Tom Matte	Division of Environmental Hazards and Health Effects, CEHIC
Howard Ory	Information Resources and Management Office
Barbara Panter	Division of Surveillance and Epidemiology, EPO
Gibb Parrish	Surveillance and Information Systems Section, CEHIC
Dave Sanders	Division of Immunizations, CPS
Lee Sanderson	Behavioral Risk Factor Surveillance Branch, CCDPHP
Ruth Slade	Division of Surveillance and Epidemiology, EPO
Joe Sniezek	Division of Environmental Hazards and Health Effects, CEHIC
Steve Teutsch	Division of Surveillance and Epidemiology, EPO
Steve Thacker	Director, Epidemiology Program Office
Wanda Tillman	Division of Surveillance and Epidemiology, EPO
Peg Tipple	Influenza Branch, Division of Viral and Rickettsial Diseases, CID
Jay Wenger	Meningitis and Special Pathogens Branch, Division of Bacterial Diseases, CID
Melinda Wharton	Division of Surveillance and Epidemiology, EPO

## APPENDIX A.

### PERSONS CONTACTED FOR THIS STUDY (CONT.)

#### II. IN STATES

<b>Minnesota</b>	
Teresa Jacques	Acute Disease Epidemiology Section
Michael Osterholm	State Epidemiologist, Acute Disease Epidemiology Section
Karen White	Acute Disease Epidemiology Section
<b>Missouri</b>	
Mahree Bright	Division of Epidemiology
Denny Donnell	State Epidemiologist, Division of Epidemiology
Peggy Fischer	Division of Epidemiology
<b>New Mexico</b>	
Margaret Gallaher	Office of Epidemiology
Judy Knott	Public Health Nurse, Albuquerque
Cathy Powers	Public Health Nurse, Albuquerque
Mack Sewell	State Epidemiologist, Office of Epidemiology
Martha Tanuz	Nurse Epidemiologist, Office of Epidemiology
<b>New York</b>	
Guthrie Birkhead	Bureau of Communicable Disease Control
Hwa-Gan Chang	Bureau of Communicable Disease Control
John Grabau	Bureau of Communicable Disease Control
Forrest Mance	Bureau of Communicable Disease Control
Dale Morse	State Epidemiologist, Bureau of Communicable Disease Control

# APPENDIX A.

## PERSONS CONTACTED FOR THIS STUDY (CONT.)

<b>Puerto Rico</b>	
Jose Pablo Aponte	Division of Epidemiology
Norma Diaz	Division of Epidemiology
Theresa Diaz	Division of Epidemiology
Daniel Jimenez	Division of Epidemiology
Edna Ponce	Division of Epidemiology
Nilka Rios	Division of Epidemiology
John Rullan	Territorial Epidemiologist, Division of Epidemiology
Anabel Santiago	Division of Epidemiology
Justina Villanueva	Division of Epidemiology
<b>South Dakota</b>	
Linda Schaefer	Division of Public Health, Communicable Disease Control
Kenneth Senger	Director, Division of Public Health
LaJean Volmer	Division of Public Health, Communicable Disease Control
Linda Zeller	Division of Public Health, Communicable Disease Control

**APPENDIX B**

**INTERVIEW GUIDES AND STATE PROTOCOL**

**MASTER LIST OF QUESTIONS TO BE  
CONSIDERED IN THE STUDY AT EPO:**

- What is the current method used to upload, store, access and maintain NETSS data at CDC? Is this in the process of modification at the present time? How are data disseminated back to States and in what time frame?
- Can the present method for collecting NETSS data handle data transmitted from multiple reporting sites within a State (e.g. STD clinics) directly to CDC?
- What differences are there in NETSS and MMWR data bases? How are they integrated at the present time for publication in weekly and annual MMWR reports?
- What quality control measures are used for incoming NETSS data? What checks are run on weekly and annual MMWR data? What role do States play in this process?
- What analyses are presently run on NETSS data at CDC? How are these disseminated? Are analyses run which are not published in the MMWR? Why? Do you know of any feedback from the States regarding analyses that they would like to have but do not at the present time?
- Can the NETSS or the NETSS/MMWR database do special reports and support State and CDC Program requests for information and assistance? Is there a potential for them to do this in the future?
- What plans are there to increase the number of CDC surveillance systems that are using NETSS? What is the status of these at the present time? Which ones are operational and which are in planning stages?
- Are there limitations to the use of extended NETSS records by CDC surveillance systems? Are there kinds of surveillance data for which NETSS is not appropriate?
- Can the NETSS record be programmed to incorporate variable and expanded record lengths to include disease-specific data? Are there any significant limitations on what can be incorporated?
- How is progress in upgrading State systems to Epi Info, Version 5? Have any problems been encountered?
- Has the expanded record format been implemented? In how many States? How is it going?

- Have the **CDC/CSTE** case definitions for cases to be reported through NETSS been published in the MMWR? Are data on this variable being submitted at the present time?
- What are the goals of the EPO workshops for State surveillance coordinators and for State epidemiologists? Are there any problems with the location, i.e. can state staff afford to attend?
- What is the status of the microcomputer based Epidemiologic Workstation? What features will it have and how will it work? What kind of hardware and staff resources will be needed for State and local health departments to install and use it?
- Do you have any opinion about the effect of planned enhancements to NETSS on reporting burden in the States? Do you think that any of the new features are likely to be a problem in the States in terms of start-up or ongoing operation?
- Do you know of concerns about confidentiality or privacy of surveillance data which have been exoressed within CDC? What technical methods are being considered to provide access to data at CDC and in States while providing acceptable protection against unauthorized access.
- What are **IRMO** plans to enhance CDC communications and data dissemination capabilities for expansion of NETSS. What is the time frame for this? Can it be done in time to support current or planned NETSS developments?
- What are the limitations of current transmission mechanism to NETSS using DIALCOM? What progress has been made to establishing direct transmission with states? Why does this seem to be a problem?
- What alternatives are there to use of the CDC communications facility for dissemination of NETSS data?

#### IN OTHER CDC PROGRAMS

- What kind of surveillance is maintained by your department? Who submits data to you and in what form? What analyses do you do? How are the data used and disseminated?
- Do you have concerns about centralized collection, processing and dissemination of surveillance data? What are they?
- Have you considered using an expanded NETSS record to obtain surveillance data from States? Have you yet reached a decision on this issue? What factors affect (affected) your decision? Was any factor especially decisive?

- Are you using some other electronic surveillance system to obtain surveillance data? What would be the implications of changing from one electronic system to another for your program? For those who provide you with data?
- Have you heard complaints about reporting burden from those who submit surveillance data? If so, how can this be addressed?
- Do you have concerns about confidentiality or privacy of surveillance data? How is this handled at the present time? Is there some reason that computer security measures will not be adequate? Do you have concerns about access to data via online bulletin boards such as WONDER?
- What hardware and software are routinely used in your program office?
- Do most of your staff feel comfortable with computers? What functions are routinely done by computer? Do program staff routinely use the CDC mainframe in doing their jobs? How about PCs? What PC analysis software is used?
- Are you familiar with **Epi Info**? Which version? How do you use it?

#### IN STATES

- Has your state had any problems with the implementation of NETSS? What were they? Were they start-up problems or have they been ongoing? How has CDC and/or EPO responded to these problems?
- Does your department use Epi Info for data entry and processing? How do you use it to support your job? What version do you use? How is version **5**? Any problems?
- Have you adopted the **CDC/CSTE** case definitions for cases to be reported through NETSS? Who is responsible for identifying cases using these case definitions? How is this checked?
- Do you have concerns about data control and confidentiality of case-level surveillance data leaving the State? What are they? Do you have any suggestions about how these concerns could be met?
- CDC is contemplating greater standardization in NETSS in contrast with the present practice of tailoring the system for each State. What do you think of this idea? Do you have special surveillance requirements or problems that demand special treatment?
- Are there any surveillance tasks which represent a notable reporting burden for your staff? What is it and why is it such a hassle? What would help alleviate this burden?

- What surveillance data do you presently submit to CDC on hard copy? is trying to expand the NETSS record to provide for transmission of all surveillance data to a central source at CDC regardless of its final destination at CDC? Would this be helpful to you for these data? Why or why not?
- What kinds of data quality control measures are applied to **surveillanc** data before it is sent to CDC? Does CDC contribute to identifying errors a 4d inconsistencies in **your surveillance data**? How?
- Are you receiving feedback on your surveillance data in a form in which you can use it? Is it timely? What kinds of dissemination do you need or want from CDC?
- CDC proposes to provide online analysis capability of the NETSS data base using WONDER, a computer bulletin board to be operated on the CDC computer. Would you use such a facility? How?
- What data analyses do you perform for your own purposes? Do you do analyses because CDC requests them? What are they? Has the installation of NETSS and Epi Info helped you produce the analyses that you need for your own purposes?
- What kinds of hardware and software are used in the health department for the management, processing and transmission of surveillance data? Are there plans to upgrade hardware capability? How firm are these plans and when will they be implemented?
- What electronic communications capabilities are presently available in between the State and local health departments? Between the State and CDC? Are there plans to expand this capabilities in the future? How firm are these plans?
- What kind of training and technical support is available to members of the State health department who are implementing computer management of surveillance data? Who provides training and technical support to your staff? Is training a routine part of the jobs of relevant staff, or must staff use initiative to identify and enter training programs? What training support do you want or need?
- How adequate have you found the technical support provided by EPO in your implementation of NETSS and Epi Info?
- Have you or any of your staff attended EPO workshops for State surveillance coordinators and for State epidemiologists? Have these been useful? Why or why not?
- What is the status of the computerization of notifiable disease data in local or regional health departments? What hardware and staff are available in local areas and what is the variability in local computer capability?

- Is it possible for the State health department or some other State agency to provide training and technical support to local health departments willing to electronically transmit infectious disease data to the State? What support would the state need to provide training at the local level? Would this be acceptable if provided by CDC or would it be best if CDC provided support to States which could train local staff?
- If local computing capacity were available, would it be feasible to transmitting data from multiple reporting sites within a State (i.e. STD clinics) directly to the CDC computer? What controls on such transmissions would be necessary for this to be acceptable to the State health department?

## EXECUTIVE SUMMARY

**TITLE.** An Evaluation of the National Electronic Telecommunications System for Surveillance (NETSS)

**CONTRACT NUMBER:** 200-88-0642

**SPONSOR:** Centers for Disease Control  
1600 Clifton Road  
Atlanta, Georgia 30333

**CONTRACTOR'S NAME**  
**AND ADDRESS:** Battelle Memorial Institute  
2101 Wilson Boulevard, Suite 800  
Arlington, Virginia 22201

[Pages 1-2] 1. Statement of the Problem

The purpose of this evaluation study was to assess the effectiveness of the National Electronic Telecommunications System for Surveillance (NETSS) in facilitating the movement of surveillance data between state health departments and the Centers for Disease Control. The study examined the fit of present and planned NETSS developments with trends in the computerization of the public health assessment function at CDC and in states and territories.

NETSS is a system of computerized record formats which is presently used to transmit National Notifiable Diseases Surveillance System (NNDSS) data between CDC and the offices of state epidemiologists. Users of NETSS employ a variety of software systems to compile data from the state's own data management system and transmit them to CDC using an electronic mail facility or direct transfer. Data are uploaded to the NNDSS database on the CDC mainframe which support production of tables and graphics in the weekly MMWR and in annual NNDSS summaries.

NETSS began in 1984 with the inception of the Epidemiologic Surveillance Project (ESP) in six states. The prime

objective of the ESP project was to assess the feasibility of replacing weekly telephone reporting of general morbidity data from states to CDC with transmission of case-level data over an interactive computer linkage. In October 1989, all 50 states, Washington, D.C., New York City and Puerto Rico were transmitting data to CDC over the system. In the same year, ESP was renamed the National Electronic Telecommunications System for Surveillance (**NETSS**) to mark its emergence as a national system for routine collection of surveillance data from states and territories.

While the evaluation reported here was the first one to be conducted of NETSS itself, EPO has commissioned two related evaluations in recent years. A 1986 evaluation of ESP by Preventive Medicine Associates in 1986 noted the diversity of the process of computerization in the states. An evaluation of systems of infectious disease surveillance maintained by CDC completed by Battelle in 1990 found that state epidemiology staff generally felt that NETSS worked well and had at least the potential to streamline surveillance and reduce reporting burden in their departments. However, problems have persisted in timely data flow between states and the CDC mainframe environment. In addition, new developments to expand the capacity of NETSS to collect program-specific data, to introduce remote entry of data by local health agencies, and to provide access to new means of data analysis and dissemination are emerging at CDC and in states and territories.

In early 1990, The CDC Surveillance Coordinating Group (**SCG**) appointed the Subcommittee on Electronic Systems for Public Health Surveillance (hereafter referred to as the SCG subcommittee) to consider methods for improving the compatibility of surveillance systems designed by CDC and to develop policy recommendations for improving CDC support of the public health

assessment function in the future. Recommendations were approved in March, 1991. Strategies for the technical implementation of these recommendations have been developed by the SCG.

Page 11

## **II. Evaluative Objectives**

This evaluation sought to build on these previous studies. Its overall goal was to support CDC program planning for future development of NETSS and other electronic surveillance initiatives at CDC and in states. The evaluation had three objectives:

- To see how the NETSS and related EPO activities operate at the present time to meet current surveillance needs at CDC and in the states,
- To see how planned enhancements to NETSS and other EPO activities will meet present and future surveillance needs at CDC and in the states, and
- To identify state and CDC needs which are not met by current or planned activities.

Recommendations for future support of surveillance were based on the findings of investigations in the EPO, other programs at CDC and in six states and territories chosen for this project.

[Pages 10-I 11

## **III. Methodology**

**This** was a case-study evaluation based on interviews with staff at CDC and in six state and territorial health departments. The study had three components: 1) establishing a baseline description of NETSS in interviews with EPO and **IRMO** staff who worked on its development, 2) investigating the present and future role of NETSS in meeting surveillance needs of other **C/I/Os** at CDC in interviews with CDC staff managing surveillance systems, and 3) examining the present and future role of NETSS in state and territorial health departments.

The primary source of data, both at CDC and in states, was interview data from public health staff who are users and operators of NETSS or who are otherwise involved in surveillance. Interview data were supported with written materials and direct observation of the operation of NETSS. A written protocol of interview questions was used in all state/territorial interviews to ensure that data from all sites were comparable and analyzable. Interviews, written documentation and observational data were compiled into reports of activities in programs and/or states. These reports were reviewed for accuracy and completeness by persons contributing data to their development. Analysis was performed with a comparative methodology in which conclusions were drawn on the basis of differences and similarities in data on key issues across staff from varying agencies, positions and states/territories.

#### **IV. Major Findings and Recommendations**

[Pages 2-91

Findings and recommendations were structured around six key topic areas which emerged in the development of this study and in the SCG subcommittee recommendations for CDC support of electronic surveillance. These are: **1)** standardization/customization of software, **2)** training and technical support, **3)** software development, **4)** telecommunications, **5)** data exchange and connectivity, and **6)** dissemination. In addition, we address developmental priorities for NETSS, and inventory the resources available in states and territories to accommodate planned enhancements of the system.

[Pages 65-661

Standardization and customization. EPO would like to move' to a more standard version of NETSS to improve support and make system upgrades more timely. The strategy followed by EPO since the inception of NETSS has been to customize the installation of the system for each user in states and territories. Customized

development of NETSS requires customized updating and operational support, imposing a limitation on the expansion of the system to a larger number of locations. In addition, states must often wait a long time for on-site upgrades of Epi Info applications.

State and territorial staff interviewed for this project did not reject the idea of a standard implementation of NETSS. Persons interviewed generally understand the need for a standard format to guarantee comparability of data from different reporting jurisdictions and to permit better quality control. Some epidemiologists supported the idea of greater standardization if it would lead to better technical support.

[Pages 74-761

Recommendations concerning standardization and customization:

- CDC should proceed with development of a more standard NETSS package. Development should include documenting the system as it is presently implemented in states and territories, defining and communicating the standard to be adopted and implementing the standard incrementally with short-term benchmarks.

[Pages 66-681

Training and technical support. For computer software to be used effectively, it must be properly installed, staff must be trained to use it, and there must be prompt access to someone knowledgeable about the software if something goes wrong. There are two models of the relationship of support to the adoption and use of software. One is to build specialized applications which maximize the capacity of the software to perform specific functions, but which require intensive technical support. A second model is to write a software package that is simple enough for the user to install and operate without intensive technical support and training. EPO has followed the first strategy with NETSS and EPI Info. EPO staff visit states and territories to install and upgrade Epi Info and provide ongoing telephone and Carbon Copy support.

States and territories place heavy reliance on EPO for technical support of start-up, ongoing operation and staff training for NETSS and Epi info. This situation is unlikely to change in the near future. State epidemiology offices do not always have access to staff with personal computer expertise. However, there is evidence that states and territories are developing their own expertise in personal computer software and in Epi Info.

[Pages 76-8 11

Recommendations concerning training and technical support:

- Providing a disk with an installation program and no further assistance or explanation is not adequate support for the kinds of major revisions envisioned by EPO.
- EPO should be prepared to provide assistance in installing and using software sent directly from **IRMO** to states if no other support is available.
- EPO should continue its on-site installation of Epi Info and NETSS in states and territories for the immediate future but should focus on capacity-building in states as a long-range goal.
- Training and technical support should rely less on “turnkey” systems and seek to educate the states on the system they are using.
- CDC should create a computer development internship to provide long-term computer expertise to states with special needs.
- CDC training in Epi Info should focus on training trainers to provide support to local health departments in the states.
- CDC should promote the teaching and use of Epi Info software in Schools of Public Health and other forums.
- CDC should support development of a mechanism for exchange of NETSS and Epi Info support among states.

[Pages 68-69]

Software development. The development of Epi Info, originally developed to support epidemic investigations and later applied to the ESP project and NETSS, has had dramatic effect on the direction of computerization of infectious disease surveillance in states and territories. In several states and territories visited for this project, Epi Info has changed the way in which surveillance is done by improving the access of health department staff to their own data. States can now analyze more data and they can do it much faster than was once possible.

Epi Info seems to be establishing itself as a standard for the management and rapid analysis of surveillance data. Epi Info was used for routine management and analyses of notifiable disease information in all the health departments visited. Even the non-Epi Info states chosen for this study, use it for some purposes. It is especially significant that it is used in all observed cases of local data entry. An increase in this practice may push states further toward an Epi Info standard.

[Pages 8 I-83]

Recommendations concerning software development:

- New or enhanced software developments should not be introduced into states and territories until both the software itself and specific applications have been well tested.
- Resources should not be devoted to competing with commercially available software that may already be in use in the states.
- Modifications to Epi Info to accommodate new operating systems and environments should be considered on a case by case basis.
- A fully LAN-Compatible EPI Info should be given a high priority.
- Epi Info enhancements should serve data management needs or specific public health objectives.

[Page 691

Telecommunications. The feasibility of direct transmission of data to CDC, timely transmission summaries and rapid dissemination of the data depend on development of a two-way capability in WONDER by IRMO. The SCG recommendations call for creation of a telecommunications “gateway” between reporting sources and the CDC communications networks via the CDC mainframe. This gateway will be developed and supported by **IRMO** with assistance from a technical advisory group. This proposal is responsive to requests by the Council of State and Territorial Epidemiologist (**CSTE**) for a single source at CDC for all disease reporting. Because so much of planned NETSS activity depends on this capability, it is essential that EPO and **IRMO** work closely together as this development proceeds.

[Pages 83-841

Recommendations concerning telecommunications:

- CDC should designate a team including representation from EPO and **IRMO** to oversee development of a telecommunications gateway between the CDC mainframe and remote users of NETSS.
- The **EPO/IRMO** team should review and update the schedule and milestones for the telecommunications gateway at least twice a year.
- There should be a joint beta test of NETSS and WONDER as soon as possible and this should be repeated at strategic points in the development of both systems.
- EPO should actively assist **IRMO** in developing the “gateway” to the CDC mainframe computer by providing technical assistance to states in the use of PC WONDER.

[Pages 70-721

Data exchange/connectivity. A network of linked computers in which all public health professionals can potentially have access to all others would provide timely access to public health data in a way that realizes the maximum potential of computers and telecommunications to make the right data instantly available at

the location where it is needed. Such a network will become a possibility with development of the CDC telecommunications “gateway”. However, development of such a network implies linkages of many kinds of reporting units including CDC programs, multiple agencies in state and local health departments, hospitals, clinics, providers, and any other involved groups.

Distributed data entry from district and local health departments is a high priority in states and territories. From the perspective of states, distributed data entry not only reduces data entry burden by decreasing the number of cases entered by any one agency, but will improve data quality by moving error resolution closer to the source of information. There is a size effect in the move toward distributed data entry. Small states have no interest in it and plan to retain control of data entry at the state level.

[Pages 85-861

Recommendations concerning data exchange/connectivity:

- Caution should be exercised in enlisting other CDC programs to adopt NETSS for their transmission. Agreements between program offices and EPO should be clearly defined and enumerated.
- CDC initiatives to encourage distributed data entry of infectious disease data in local health departments should be continued.
- Inter-state and direct intra-state connectivity is a low priority for states and territories visited in this evaluation, and should not be a high priority for CDC.

[Pages 72-731

Dissemination. Timely dissemination of data is a high priority for CDC and for public health officials involved in disease control programs. Surveillance data are “data for action” as well as for documentation of health events. If they are not available in time for action, their usefulness is reduced. Access to national

surveillance data in a timely fashion has been a deficiency of many CDC surveillance systems for infectious disease. Computers have the potential to improve the performance of CDC in providing rapid turnaround of surveillance data.

The usefulness of infectious disease surveillance in states is improved by rapid collection and dissemination made possible by computer transmission of data because it improves the access of states to their own data. However, the advantages to states of rapid turnaround of national surveillance reports are less clear. National surveillance is neither timely enough to support the control of infectious disease at the state level nor is it needed for this purpose. The rapid turnaround of data may be less important than is better quality control and analysis capability.

[Page 861

Recommendations concerning dissemination:

- The most useful rapid turnaround summaries for states and territories would be line-listed reports of transmissions which would permit ongoing reconciliation of data sets.

[Pages 86-891

Developmental priorities

There are a host of upgrades and enhancements to NETSS in various stages of planning and implementation at EPO which need to be prioritized. We recommend the following priorities for immediate development.

- First priority: The ADABASE/NATURAL re-write of the NETSS system at CDC should be given highest priority in terms of resources and in terms of a focus for planning.
- Second priority: Detailed documentation of the NETSS system as it is presently installed in the states and territories should be **EPO's** second priority .
- Third priority: The degree to which state installations can be standardized, from the use of common Epi Info programming to a standard reporting form, should be determined and a standard developed.

- Fourth priority: The standard implementation of NETSS then should be introduced to states and territories.
- Fifth priority: EPO should do everything possible to promote an effective telecommunications “gateway” between states and CDC.
- Sixth priority: CDC should implement measures to provide adequate technical support to states seeking distributed data entry at the county level.
- Ongoing: Software development priorities should be established on the basis of compatibility with the standard NETSS implementation and should not compete for resources with activities needed to maintain surveillance.

[Page 731

#### Resource **Needs in States and Territories**

States and territories have or will soon have the hardware necessary to support present and planned enhancements to NETSS. All of the state and territorial health departments visited have computers with at least a 286 processor. Three of them have 386 processors. Disk space is adequate to run the Epi Info software, although storage of year-to-date NETSS data may become a problem in states with a very large number of cases. The data transmission hardware and software needed for a direct link to the CDC mainframe is already **present** in state and is used to send NETSS transmission to DIALCOM.